Sparse Representations & Compressed Sensing with application to the problem of Direction-of-Arrival estimation

Aris Gretsistas

Thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

School of Electronic Engineering and Computer Science
Queen Mary University of London

April, 2013

I declare, as the author of this thesis, that the material contained within this thesis is to the best of my knowledge, original work, and all references are cited accordingly. Information from it may be freely used with acknowledgement.

Aris Gretsistas

Abstract

The significance of sparse representations has been highlighted in numerous signal processing applications ranging from denoising to source separation and the emerging field of compressed sensing has provided new theoretical insights into the problem of inverse systems with sparsity constraints.

In this thesis, these advances are exploited in order to tackle the problem of direction-of-arrival (DOA) estimation in sensor arrays. Assuming spatial sparsity e.g. few sources impinging on the array, the problem of DOA estimation is formulated as a sparse representation problem in an overcomplete basis. The resulting inverse problem can be solved using typical sparse recovery methods based on convex optimization i.e. ℓ_1 minimization. However, in this work a suite of novel sparse recovery algorithms is initially developed, which reduce the computational cost and yield approximate solutions. Moreover, the proposed algorithms of Polytope Faces Pursuits (PFP) allow for the induction of structured sparsity models on the signal of interest, which can be quite beneficial when dealing with multi-channel data acquired by sensor arrays, as it further reduces the complexity and provides performance gain under certain conditions.

Regarding the DOA estimation problem, experimental results demonstrate that the proposed methods outperform popular subspace based methods such as the multiple signal classification (MUSIC) algorithm in the case of rank-deficient data (e.g. presence of highly correlated sources or limited amount of data) for both narrowband and wideband sources. In the wideband scenario, they can also suppress the undesirable effects of spatial aliasing.

However, DOA estimation with sparsity constraints has its limitations. The compressed sensing requirement of incoherent dictionaries for robust recovery sets limits to the resolution capabilities of the proposed method. On the other hand, the unknown parameters are continuous and therefore if the true DOAs do not belong to the predefined discrete set of potential locations the algorithms' performance will degrade due to errors caused by mismatches. To overcome this limitation, an iterative alternating descent algorithm for the problem of off-grid DOA estimation is proposed that alternates between sparse recovery and dictionary update estimates. Simulations clearly illustrate the performance gain of the algorithm over the conventional sparsity approach and other existing off-grid DOA estimation algorithms.

στους γονείς μου

Acknowledgements

My deepest gratitude goes to my supervisor Prof. Mark Plumbley, who made this PhD study an astonishing and rewarding experience with his invaluable assistance, support and guidance. Without him this research project would not have been possible. Also, I must acknowledge my secondary supervisor Dr Josh Reiss for his constructive feedback on many aspects of this work.

Many thanks to all past and present members of C4DM at QMUL, who have made it an amazing place to work providing their knowledge, friendship and entertainment. In particular, a special mention goes to Ken O'Hanlon, Ivan Damnjanovic, Dan Stowel, Boris Mailhé, Daniele Barchiesi, Andrew Nesbit, Maria Jafari, Nicolae Cleju, Luís Figueira, Hidehisa Nagano, Hiromasa Fujihara, Samer Abdallah, Constantin Barabasa, Fabio Hedayioglu, Tim Murray Browne, Andrew Robertson, Steve Welburn, Adam Stark and Dimitrios Giannoulis for their continuous assistance in many theoretical and practical issues encountered throughout the duration of this PhD study.

I am grateful to several people involved in the "SMALL" project. Thank you, Rémi Gribonval, Mike Davies, Michael Elad, Pierre Vandergheynst, Simon Arberet, Valentin Emiya, Sangnam Nam, Alexis Benichoux, Mehrdad Yaghoobi, Mohammad Golbabaee and Nancy Bertin for all your fruitful comments and all your motivating discussions.

Above all, I thank my parents for their infinite support throughout everything. Last but not least, I would like to thank Silvia Bologna for her support and encouragement, when I have needed it the most.

This research was supported by the EPSRC Leadership Fellowship EP/G007144/1 and the EU FET-Open Project FP7-ICT-225913 "SMALL".

Contents

List of Figures

List of Tables

List of Abbreviations

ACG-PFP	Approximate Conjugate Gradient Polytope Faces Pursuit
AWGN	Additive White Gaussian Noise
BOMP	Block Orthogonal Matching Pursuit
BP	Basis Pursuit
BPDN	Basis Pursuit Denoise
BSS	Blind Source Separation
CG	Conjugate Gradient
CG-PFP	Conjugate Gradient Polytope Faces Pursuit
CG-StPFP	Stagewise Conjugate Gradient Polytope Faces Pursuit
CRB	Cramér-Rao Bound
CS	Compressed Sensing
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DL	Dictionary Learning
DML	Deterministic Maximum Likelihood
DOA	Direction-Of-Arrival
DWT	Discrete Wavelet Transform
ERC	Exact Recovery Condition
FFT	Fast Fourier Transform
G-SOMP	Group Simultaneous Orthogonal Matching Pursuit
GOMP	Group Orthogonal Matching Pursuit
GPFP	Group Polytope Faces Pursuit
GPSR	Gradient Projections for Sparse Reconstruction
KKT	Karush-Kuhn-Tucker
LARS	Least Angle Regression
Lasso	Least Absolute Shrinkage and Selection Operator
LP	Linear Programming
LS	Least Squares
ML	Maximum Likelihood
MMV	Multiple Measurement Vector
MP	Matching Pursuit
MUSIC	MUltiple SIgnal Classification
MVDR	Minimum Variance Distortionless Response
NLA	Non-uniform Linear Array
OMP	Orthogonal Matching Pursuit

PFP	Polytope Faces Pursuit
RA-ORMP	Rank Aware Orthogonal Regularized Matching Pursuit
RA-PFP	Rank Aware MMV-Polytope Faces Pursuit
RIP	Restricted Isometry Property
RMSE	Root-Mean-Square Error
SBI	Sparse Bayesian Inference
SDP	Semi-Definite Programming
SML	Stochastic Maximum Likelihood
SMV	Single Measurement Vector
SNR	Signal-to-Noise-Ratio
SOCP	Second Order Cone Programming
SRTLS	Sparse Regularized Total Least Squares
STFT	Short-Time Fourier Transform
StOMP	Stagewise Orthogonal Matching Pursuit
SVD	Singular Value Decomposition
SWOMP	Stagewise Weak Orthogonal Matching Pursuit
TLS	Total Least Squares
ULA	Uniform Linear Array

List of Symbols

$(\cdot)^T$	transpose of vector or matrix
$(\cdot)^H$	Hermitian transpose of vector or matrix
$(\cdot)^{-1}$	inverse of a matrix
$(\cdot)^\dagger$	pseudo-inverse of a matrix
$\lvert\cdot\rvert$	absolute value of a scalar
$\lVert\cdot\rVert_p$	ℓ_p norm of a vector or a matrix
$\lVert\cdot\rVert_{p,q}$	mixed $\ell_{p,q}$ norm of a vector or a matrix
$\langle\cdot,\cdot\rangle$	inner product between two vectors
∇	gradient operator
∂	partial differential operator
$\odot$	Schur-Hadamard matrix product
$\otimes$	Kronecker matrix product
$\bigcup$	union of sets
$\mathbb{C}$	set of complex numbers
$\mathbb{R}$	set of real numbers
$\mathbb{P}(\cdot)$	probability of an event
$\mathcal{M}$	nonlinear manifold
$\mathcal{O}$	Big O notation
$\lambda_{\max}(\cdot)$	maximum eigenvalue of a matrix
μ	mutual coherence of a matrix
σ^2	noise variance
$\mathrm{conv}\{\cdot\}$	convex hull of a set of vertices
$\mathrm{diag}\{\cdot\}$	diagonal matrix
$\emptyset$	empty set
$\inf\{\cdot\}$	infimum of a subset
$\mathrm{null}(\cdot)$	nullspace of matrix
$\mathrm{orth}(\cdot)$	orthogonalization of columns of matrix
$\mathrm{rank}(\cdot)$	rank of a matrix
$\mathrm{spark}(\cdot)$	spark of a matrix
$\mathrm{sinc}(\cdot)$	sinc function
$\mathrm{sgn}(\cdot)$	sign or signum function
$\sup\{\cdot\}$	supremum of a subset
$\mathrm{supp}(\cdot)$	support set of a vector
$\mathrm{tr}(\cdot)$	trace of a matrix
$\mathrm{E}\{\cdot\}$	statistical expectation operator

$\mathcal{I}$	set of indices
$I(\cdot)$	indicator function
K	sparsity level (cardinality of the true support set)
$\mathcal{N}(\cdot)$	normal distribution
S ($\hat{S}$)	true column support set of a (doubled) matrix
$\bar{S}$	complement of the true support set
$\mathbf{0}$	vector of zeros
$\mathbf{1}$	vector of ones
$\mathbf{I}$	identity matrix
$\mathbf{a}_i$	i-th column vector of matrix $\mathbf{A}$
$\mathbf{A}_S$	the columns of matrix $\mathbf{A}$ corresponding to the true support set S
$P_{\mathbf{A}}$	projection onto the space spanned by $\mathbf{A}$
$P_{\mathbf{A}}^{\perp}$	projection onto the nullspace of $\mathbf{A}$

Chapter 1

Introduction

One of the main challenges in many signal processing applications is to furnish compact and effective representations of high dimensional data that can retain the main structures inherent in the signal of interest. These compact representations are commonly referred to as sparse representations and they are concerned with the construction of signal models that involve the linear combination of a few elementary components (known as atoms) drawn from a large set of vectors (known as a dictionary). The field of sparse representations has rapidly evolved to become a very large scientific research area, with widespread applications, ranging from data acquisition and signal recovery to denoising and source separation. As an outcome, a significant amount of study has been conducted in order to give answers to questions such as how to find "good" bases or overcomplete dictionaries in terms of adaptability for a given class of signals. This extensive study has led to some interesting theoretical developments, and the birth of related fields such as compressed sensing. Compressed sensing provides a new revolutionary sampling scheme, which based on the sparsity assumption goes beyond conventional sampling principles, posing a challenge to traditional signal processing compression fundamentals. According to this scheme, compression is achieved by subsampling at very low rates. The underlying dimensionality reduction is accomplished by deploying random projections, which can have great importance from a practical perspective. When the sparsity assumption holds, the undersampled data can be sufficient for the perfect recovery of the original signal. In this sense, the framework of compressed sensing promises unification of the

concepts of sampling and compression.

Notwithstanding that much progress has been made in the field of sparse representations and compressed sensing, along with their related applications, it is generally believed that the aforementioned frameworks have not yet reached their full potential, since many issues in practical applications either have not yet been completely realised or have remained unexplored. To the best of the author's knowledge, among various applications direction-of-arrival (DOA) estimation and source localization is a major field that the exploitation of sparsity has not been fully utilized.

The main objective in DOA estimation is to find the location of the sources impinging on an array of sensors. In order to do this, the spatial field is first sampled and then the obtained measurements are processed in order to retrieve information about the location of the sources. This information is of crucial importance for many signal processing tasks. For instance, in an acoustic environmental setting, accurate localization of the sound sources could lead to audio enhancement or improved source separation.

1.1 Motivations

The groundbreaking work of Malioutov et al. [3] has shown that the source localization problem can be formulated as an inverse sparse recovery problem when the assumption of spatial sparsity holds, implying that there are only a few sources to be localized. Unlike the classic array processing methods that treat the problem of DOA estimation as a parameter estimation problem, the sparsity based approach has proven to be quite successful in scenarios where the amount of samples is not sufficient or when the incoming sources are highly correlated. However, the sparsity based approach involves the inversion of a non-trivial underdetermined system of linear equations through the appropriate regularization. Although recent advancements in sparse regularization have shown that the optimization task can be tackled using principled methods that perform global optimization, their convergence might be slow for practical applications dealing with large scale data, including DOA estimation. On the other hand, heuristic based greedy approaches might provide faster convergence with the price of lower precision solutions. This is due to the fact that greedy methods are iterative in nature performing a local optimization at each step instead of minimizing a global criterion. Apart from the

sparse regularization, little effort has been made to incorporate further structure to the problem.

All these observations give rise to questions such as: How one can build more efficient application oriented methods that could also incorporate additional structures? Could the compressed sensing framework provide further meaningful insights to the problem of DOA estimation? Is there any link between the classic array processing methods and the sparsity based ones?

In this work, all these questions are addressed and for this reason, the problem of DOA estimation is thoroughly examined from a sparse representation point of view. As a result, this work contributes on building efficient algorithms that incorporate additional structures, without overlooking the theoretical aspects of the problem.

1.2 Aim

The aim of this thesis is to investigate sparse recovery methods and their application to the problem of DOA estimation. To do this, the focus is initially drawn on the sparse representation problem and an attempt is made to develop novel algorithms of general use that leverage the benefits of the principled regularization approaches with the advantages of the greedy approaches. Among the main objectives of this thesis is to provide algorithms suitable for large scale recovery problems that can also incorporate additional structure that the signals of interest might exhibit. The goal then is to show how these developed structures can be applied to the problem of DOA estimation and source localization. The identification of some limitations of the specific framework leads to the proposal of novel techniques that serve as improvements to the original sparsity based approach.

1.3 Thesis overview

Chapter 2 contains the background and existing research with which this thesis is mainly concerned. First of all, the problem of DOA estimation in sensor arrays is formally introduced. The signal model for the scenario of far-field narrowband propagation is built and the main array processing methods related to this work are presented. After identifying the link between the array processing methods and sparsity, a thorough introduction

to the basic fundamentals and principles of sparse recovery and compressed sensing is provided.

Chapter 3 focuses on the development of fast sparse recovery algorithms, which are suitable for large scale inverse problems. More specifically, a fast implementation of the iterative method of Polytope Faces Pursuit (PFP) [1] is proposed that updates the solution vector at each iteration using the method of conjugate gradients. Stepwise and stagewise criteria are adopted in the discussion that follows and rigorous theoretical findings along with experimental results demonstrate the robustness of the proposed algorithms.

Chapter 4 investigates the problem of structured sparsity. In more detail, the group sparse recovery problem is first introduced and alternative constraints that need to be imposed on the optimization task are examined. Therefore, the conventional sparsity algorithm is extended to the group sparsity case in order to develop the Group Polytope Faces Pursuit (GPFP) algorithm. Theoretical and empirical results demonstrate the superiority of GPFP over the conventional sparsity approach in the group sparse scenario. Finally, the problem of joint sparsity over multiple measurement vectors is presented and accordingly an extension of the original algorithm is proposed. Although theoretical results fail to predict any performance gain over the conventional sparsity model, experimental results reveal its improved performance in joint sparse recovery.

Chapter 5 investigates the applicability of the developed sparse recovery algorithms to the problem of DOA estimation and source localization. After the angular space is appropriately discretized and the importance of various parameters of the problem, such as spacing and number of sensors, is highlighted, it is shown that under certain assumptions (i.e. the number of sources or the noise level is known) the developed suite of algorithms can achieve accurate DOA estimation. Experimental results confirm the advantages of the proposed approach over the classic array processing techniques. At the same time, certain modifications of the algorithm also provide enhanced performance when compared to the sparsity based work of [3]. The problem of wideband DOA estimation assuming far-field and near-field propagation is then examined. The presented work shows how limitations of the classic array processing methods, such as spatial aliasing can be overcome in this case by exploiting the underlying structured sparsity.

Chapter 6 presents a novel algorithm that tackles the problem of off-grid DOA estimation under sparsity constraints. The main limitation of the sparsity based source localization framework is the assumption that the DOAs of the incoming sources exactly match the ones contained in the discrete angular grid. Instead of constructing a finer grid, the proposed approach is an iterative algorithm, which at each step attempts to redefine the dictionary based on the Taylor approximation. Experimental results demonstrate the performance gain when the DOAs of the sources are off the predefined grid.

Chapter 7 concludes this thesis and summarizes its main contributions. Prospects for further research are also considered.

1.4 Associated publications

This thesis is partly derived from the following previously published work by the author:

- Parts of Chapter 3 were originally published at the 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2010) [4] and at the 2010 conference on Information Representation and Estimation (INSPIRE 2010) [5].

- Parts of Chapter 4 were originally presented at the 10th international conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2012) [6].

- Parts of Chapter 5 were originally presented at the 9th international conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2010) [7] and at the 2011 workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS11) [8].

- Parts of Chapter 6 were originally published at the 20th European Signal Processing Conference (EUSIPCO 2012) [9].

Chapter 2

Background

To establish the basis upon which this thesis will be developed, this chapter provides the background theory of the main research topics related to this work. First of all, the discussion focuses on the problem of direction-of-arrival (DOA) estimation using sensor arrays. The signal model is built and the main array processing approaches to spatio-temporal spectral analysis are presented. A short overview of the methods that attempt to reformulate the DOA estimation problem as a sparse recovery problem follows.

Then, the purposes and main concepts of sparse signal representations and compressed sensing are examined. The chapter closes with a brief discussion of their basic theoretical background and most popular algorithms.

2.1 Direction-of-arrival estimation

Direction-of-arrival estimation using sensor arrays has been an active research field, playing a fundamental role in many signal processing areas such as radar [10], sonar [11], seismology [12] and acoustic tracking [13]. The main objective in the DOA estimation or source localization problem is to estimate the spatial energy spectrum and therefore determine the number and location of the sources of energy. To do this, temporal and spatial information is first obtained by sampling the wave field with sensor arrays and then processed with the aim to reveal the directions of the emitting sources that form this wave field. Put another way, DOA estimation is a parameter estimation problem. Its origins date back to the 1940s, when the first attempt on spectral analysis using

spatio-temporally sampled data was conducted [14]. From then onwards, there has been ongoing research in the field of source localization with the goal of developing methods that do not only yield accurate estimates under ideal conditions, but more importantly are robust to non ideal conditions such as noisy measurements, limitations on the number of measurements, the aperture size of the array or the number of sensors.

What follows serves as a brief introduction to the problem of DOA estimation of sources that impinge on a linear array of sensors. After the formal description of the conventional array signal model, an overview of the most popular DOA estimation methods from the field of array processing is given. Finally, this section outlines that under certain assumptions a formulation of the DOA problem as an underdetermined inverse sparse recovery problem is plausible. This has also been highlighted in numerous recently emerging alternative methods [3, 15, 16, 17, 18, 19, 20].

2.1.1 Signal model & problem formulation

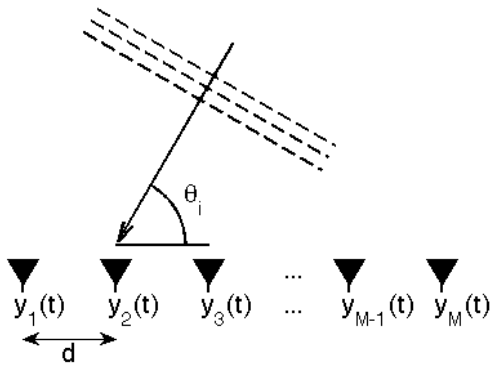


Figure 2.1: Impinging plane wave on a uniform linear array of M sensors with inter-element spacing d.

Consider a uniform linear array (ULA) of M sensors with inter-element spacing d, as illustrated in Fig. 2.1. The sensors sample spatially the wave field, which is assumed to be generated by a finite number of emitting sources. The sources are assumed to have negligible extent relative to the aperture size of the array, so that they can be modelled as point sources. The medium is considered homogeneous and therefore the propagating speed is constant. The propagating waves corresponding to the emitters are considered

either spherical or planar waves, depending on the distance between the ULA and the location of the emitting sources [3]. In the former case, which is known as the near-field case, the sources are located relatively close to the array; while in the latter case, known as the far-field propagation model, the location of the sources is far with regards to the aperture size of the array.

For simplicity without loss of generality, it is assumed that one plane wave propagating from the far-field impinges on the array from an unknown direction (Fig. 2.1). It is also assumed that the signal is narrowband (i.e. digital modulated signal with carrier frequency f_c). A narrowband source is modelled as a complex envelope (or complex bandpass signal):

$$\hat{x}(t) = x(t)e^{j\omega_c t} \tag{2.1}$$

where $\omega_c = 2\pi f_c$ is the carrier frequency and $x(t)$ is the baseband signal [21, 22]. Each sensor captures the incoming signal with a time delay. In the noiseless case, the signal received by the m-th sensor is given by:

$$y_m(t) = x(t - \tau_m)e^{j\omega_c(t-\tau_m)}. \tag{2.2}$$

The narrowband assumption implies that the spectrum of the narrowband signal is band-limited to the region:

$$|\omega_L| \leq \pi B_s \tag{2.3}$$

where $\omega_L \triangleq \omega - \omega_c$ and πB_s specifies the maximum signal bandwidth. If it happens that the bandwidth of the signal is much less than $1/\tau_m$ ($B_s\tau_m << 1$), then one can make use of the narrowband approximation, which allows to ignore the delay τ_m from the baseband signal $x(t - \tau_m) \approx x(t)$. This is because in that case the signal changes very slowly relative to the travel time across the aperture of the ULA. Taking this approximation into account, equation (2.2) becomes:

$$y_m(t) \approx x(t)e^{j\omega_c t}e^{-j\omega_c\tau_m}, \quad \text{for} \quad m = 1, 2, \ldots, M. \tag{2.4}$$

In practice, the dependence on the term $e^{j\omega_c t}$ is usually dropped. It follows that the

sensors will capture:

$$\begin{aligned} y_1(t) &= \hat{x}(t-\tau_1) \propto x(t)e^{-j\omega_c\tau_1} = x(t)e^{-j2\pi f_c\tau_1} \\ y_2(t) &= \hat{x}(t-\tau_2) \propto x(t)e^{-j\omega_c\tau_2} = x(t)e^{-j2\pi f_c\tau_2} \\ &\cdots \quad \cdots \quad \cdots \\ y_M(t) &= \hat{x}(t-\tau_M) \propto x(t)e^{-j\omega_c\tau_M} = x(t)e^{-j2\pi f_c\tau_M} \end{aligned} \tag{2.5}$$

where $\tau_m = (m-1)d\cos(\theta)/c$ if the first sensor of the array is the phase reference, c is the propagation speed and m represents the sensor index. Therefore, the sensor array output can be modelled as:

$$\mathbf{y}(t) = \begin{bmatrix} e^{-j2\pi f_c\tau_1} \\ e^{-j2\pi f_c\tau_2} \\ \cdots \\ e^{-j2\pi f_c\tau_M} \end{bmatrix} x(t) + \mathbf{n}(t) = \mathbf{a}(\theta)x(t) + \mathbf{n}(t) \tag{2.6}$$

where $\mathbf{n}(t) = [n_1(t),\ \ldots,\ n_M(t)]^T$ is the $M \times 1$ vector corresponding to the additive noise at the sensors and $\mathbf{a}(\theta)$ is the linear array response to the impinging plane wave that can be expressed as:

$$\mathbf{a}(\theta) = \left[e^{-j2\pi f_c\tau_1},\ e^{-j2\pi f_c\tau_2},\ \ldots,\ e^{-j2\pi f_c\tau_M}\right]^T \tag{2.7}$$

or:

$$\mathbf{a}(\theta) = \left[1,\ e^{-j2\pi f_c d\cos(\theta)/c},\ \ldots,\ e^{-j2\pi(M-1)f_c d\cos(\theta)/c}\right]^T \tag{2.8}$$

after the substitution of the time delays τ_m for each sensor. Equation (2.6) can be easily generalized for multiple directions of arrival corresponding to multiple propagating plane waves:

$$\mathbf{y}(t) = \sum_{j=1}^{K} \mathbf{a}(\theta_j)x_j(t) + \mathbf{n}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) + \mathbf{n}(t) \tag{2.9}$$

where

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1),\ \mathbf{a}(\theta_2),\ \ldots,\ \mathbf{a}(\theta_K)] \tag{2.10}$$

is the $M \times K$ matrix containing the array responses to all impinging plane waves,

$$\mathbf{x}(t) = [x_1(t),\ x_2(t),\ \ldots,\ x_K(t)]^T \tag{2.11}$$

is the $K \times 1$ vector that contains the K plane waves impinging on the ULA and

$$\boldsymbol{\theta} = [\theta_1,\ \theta_2,\ \ldots,\ \theta_K]^T \tag{2.12}$$

is the $K \times 1$ vector that contains the DOAs of the incoming signals.

2.1.2 Spatial aliasing

In the above discussion, a ULA geometry with inter-sensor spacing d was assumed. However, for the design of a non-degenerate array structure (which is spatial aliasing free) the spacing d should be chosen appropriately. More specifically, the phase difference should be restricted to π:

$$2\pi f_c \Delta\tau \leq \pi \tag{2.13}$$

in order to avoid the undesirable effects of spatial aliasing. This type of aliasing is identical to the problem of aliasing in time series analysis and can introduce ambiguities to the non-trivial task of DOA estimation, which may make localization impossible [21]. This places an important restriction on the geometry of the ULA. Replacing $\Delta\tau = d\cos(\theta)/c$ in equation (2.13) and after some manipulation, it yields:

$$d \leq \frac{1}{2}\frac{c}{f_c}\frac{1}{\cos(\theta)}. \tag{2.14}$$

The denominator of the right hand side of the above inequality takes its maximum value at $\theta = 2k\pi$, where $k = 0, 1, 2, \ldots$. Therefore, substituting $\cos(\theta) = 1$ and the wavelength $\lambda = c/f_c$, equation (2.14) reduces to the following inequality:

$$d \leq \lambda/2 \tag{2.15}$$

which means that the inter-sensor spacing should not exceed half the wavelength of the narrowband signal.

2.1.3 Array processing DOA estimation methods

The classical array processing methods can be divided into parametric methods, which are based on the maximum likelihood (ML) paradigm and spectral based approaches often referred to as non-parametric approaches [13]. The former include deterministic

maximum likelihood (DML) and stochastic maximum likelihood (SML), where the signal waveforms are treated as deterministic and stochastic processes respectively. After the likelihood function has been obtained, the unknown parameters corresponding to the unknown DOAs are estimated so that the likelihood function is maximized. The parametric approaches result in accurate estimates at the price of high computational complexity. Since parametric methods are not the main concern of this thesis, their formal description has been omitted, but the interested reader can find more details in [13, 21].

On the other hand, non-parametric methods are computationally attractive and can be divided into two main subcategories; the beamforming techniques and the subspace based methods. The beamforming techniques attempt to steer the array in one direction at a time and measure its output power at the specific direction. Therefore, the locations that yield the maximum power are the DOA estimates. In contrast, spectral based methods employ subspace analysis and exploit the fact that the noise subspace is orthogonal to the signal subspace.

Conventional beamformer

The conventional (or Bartlett) beamformer is probably the most classic DOA estimation technique and can be viewed as an extension of the Fourier spectral analysis [23] to accommodate sensor array data. The task of DOA estimation is accomplished by "steering" the array at different locations through the appropriate weighting (or shifting) of the waveforms captured by each sensor of the array. The beamformer output can be written as:

$$\begin{aligned} P(\boldsymbol{\theta}) &= \mathbf{w}^H \mathrm{E}\{\mathbf{y}(t)\mathbf{y}^H(t)\}\mathbf{w} \\ &= \mathbf{w}^H \mathbf{R}_\mathbf{y} \mathbf{w} \end{aligned} \tag{2.16}$$

where $\mathrm{E}\{\cdot\}$ is the expectation, $\mathbf{R}_\mathbf{y} = \mathrm{E}\{\mathbf{y}(t)\mathbf{y}^H(t)\}$ is the spatial autocorrelation matrix of the zero-mean spatially stationary random variable $\mathbf{y}(t)$ and $\mathbf{w}$ represents the unknown complex weight vector. Considering the signal model of equation (2.9), $\mathbf{R}_\mathbf{y}$ can be written

as:

$$\begin{aligned}\mathbf{R_y} &= \mathrm{E}\big\{[\mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)+\mathbf{n}(t)][\mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)+\mathbf{n}(t)]^H\big\} \\ &= \mathbf{A}(\boldsymbol{\theta})\mathrm{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\}\mathbf{A}^H(\boldsymbol{\theta}) + \mathrm{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\} \\ &= \mathbf{A}(\boldsymbol{\theta})\mathbf{R_x}\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2\mathbf{I}\end{aligned} \tag{2.17}$$

where the noise is assumed white i.e. $\mathbf{n}(t) \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ and $\mathbf{R_x}$ is the autocorrelation matrix of the sources. The problem of maximizing the beamformer power can then be formulated as:

$$\begin{aligned}\max_{\mathbf{w}} \quad & \mathbf{w}^H\mathbf{A}(\boldsymbol{\theta})\mathbf{R_x}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{w} + \sigma^2\|\mathbf{w}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{w}\|_2^2 = 1.\end{aligned} \tag{2.18}$$

When only one source is present, the solution to the above problem of equation (2.18) is:

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\theta)}{\sqrt{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}} = \frac{\mathbf{a}(\theta)}{\|\mathbf{a}(\theta)\|}. \tag{2.19}$$

For multiple sources, multiple peaks are expected in the spatial spectrum:

$$P_{BF}(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{R_y}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)} \tag{2.20}$$

and can be obtained by a one dimensional search for the K highest maxima. Nevertheless, $\mathbf{R_y}$ is considered unknown and in practical array processing applications the sample covariance matrix is used instead as an approximate estimate:

$$\hat{\mathbf{R}}_\mathbf{y} = \frac{1}{T_s}\sum_{t=1}^{T_s}\mathbf{y}(t)\mathbf{y}^H(t) \tag{2.21}$$

where T_s is number of time samples or snapshots.

The conventional beamformer, reminiscent of the classic periodogram, is known for its algorithmic simplicity. However, it has certain limitations. Most notably, it suffers from the Rayleigh resolution limit, as it cannot resolve two closely spaced sources and its performance is limited by the aperture size of the ULA. The beampattern of a ULA

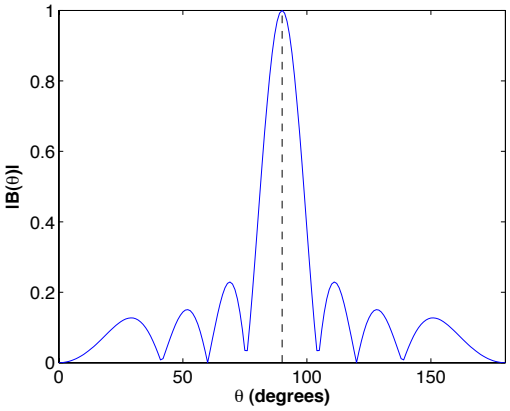


Figure 2.2: Beam pattern $|B(\theta)|$ of a uniform linear array of $M = 8$ sensors with inter-element spacing $d = \lambda/2$ for the range $\theta \in [0°, 180°]$.

of M sensors is defined as:

$$B(\theta) = \frac{1}{M}\sum_{m=1}^{M} e^{-j2\pi f_c \tau_m} = \frac{1}{M}\sum_{m=1}^{M} e^{-j2\pi f_c (m-1) d\cos(\theta)/c} \Rightarrow$$
$$B(\phi) = \frac{1}{M}\sum_{m=0}^{M-1} e^{jm\phi} = \frac{1}{M}\frac{1-e^{jM\phi}}{1-e^{j\phi}} \tag{2.22}$$

where ϕ is the electrical angle defined as $\phi = -\frac{\omega_c}{c} d\cos(\theta)$. Fig. 2.2 shows $|B(\theta)|$ in dB of a ULA of $M = 8$ sensors with half wavelength inter-element spacing and as can be seen the main lobe is centered at 90°. The Rayleigh resolution limit is defined as the distance from the central angle of the main lobe to angle corresponding to the first null of the beampattern $B(\theta)$. A null occurs when the numerator of (2.22) is zero, while its denominator is different from zero:

$$B(\phi) = 0 \Rightarrow 1 - e^{-jM\phi} = 0 \Rightarrow \cos(M\phi) = 1. \tag{2.23}$$

It follows from (2.23):

$$M\phi = 2k\pi, \quad k = 1, 2, \ldots \tag{2.24}$$

The first null occurs at $k = 1$. The corresponding value of the electrical angle $\phi = 2\pi/M$ defines the Rayleigh resolution limit. A conventional beamformer will be able to resolve two plane waves if the peak of second beam pattern lies outside the null of the first beam pattern. Consequently, for a ULA of M sensors, the conventional beamformer requires

that:

$$\phi \geq 2\pi/M. \tag{2.25}$$

For example, consider the beam pattern of Fig. 2.2. Since $M = 8$ the first null occurs at $\phi = \pi/4$. In this case, the sensor spacing is $d = \lambda/2$ and therefore it holds that $\phi = -\pi\cos(\theta)$. After some manipulation, it can be shown that $\theta \approx 104.5°$. Considering that the peak of the main lobe of the beam pattern occurs at 90°, it turns out that the Rayleigh resolution limit for this array is $\Delta\theta \approx 14.5°$. This is the minimum distance between two plane waves so that they are resolvable by a conventional beamformer. For a longer ULA of $M = 10$ sensors the Rayleigh resolution limit is $\Delta\theta \approx 11.5°$. This shows how a longer array possesses improved resolution capabilities.

Minimum Variance Distortionless (MVDR) beamformer

The MVDR beamformer proposed in [24] by Capon, also known as Capon's beamformer, provides an improved spectral estimation method that resolves the resolution limitations of conventional beamforming. MVDR replaces (2.18) with the optimization problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{R}_\mathbf{y} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{a}(\theta) = 1 \end{aligned} \tag{2.26}$$

where $\mathbf{a}(\theta)$ is defined as in equation (2.8) and $\mathbf{w}^H\mathbf{a}(\theta) = 1$ is the minimum distortion constraint. In contrast to the conventional beamformer, which attempts to maximize the output power in the "look direction" θ, the MVDR beamformer attempts to minimize the noise power and the power contributed by signals impinging on the array from other directions than θ with the constraint of unit gain in the "look direction" θ. The solution to problem (2.26) is given in [13] resulting in:

$$\mathbf{w}_{CAP} = \frac{\mathbf{R}_\mathbf{y}^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{R}_\mathbf{y}^{-1}\mathbf{a}(\theta)} \tag{2.27}$$

and the obtained spatial spectrum is:

$$P_{CAP}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{R}_\mathbf{y}^{-1}\mathbf{a}(\theta)}. \tag{2.28}$$

Once again the sample covariance matrix $\hat{\mathbf{R}}_\mathbf{y}$ can be used in practical applications.

Capon's beamformer reduces the spectral leakage caused by closely spaced sources that limits the resolution capability of the conventional beamformer. It can be viewed as an optimal beamformer and this is why it has found extensive use in practical applications [13]. Despite that, its performance is still dependent on the aperture size and the noise level.

MUSIC algorithm

Subspace based approaches to DOA estimation possess high resolution capabilities and for this reason they have been studied thoroughly. Among them multiple signal classification (MUSIC) [25, 13, 21] is a popular, powerful tool for the problem of spectral analysis and system identification. MUSIC initially obtains an estimate of the covariance matrix of the observations. This is followed by a subspace analysis, in which the covariance matrix is firstly decomposed and the space spanned by the received data is then partitioned into the signal and the noise subspace.

Consider the covariance matrix $\mathbf{R}_\mathbf{y}$ of equation (2.17). An eigenvalue decomposition of $\mathbf{R}_\mathbf{y}$ can reveal that after arranging the eigenvalues in descending order it holds:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_K > \lambda_{K+1} = \ldots = \lambda_M = \sigma^2. \tag{2.29}$$

Assuming moderate noise levels, the first K eigenvalues are much larger than the last $M-K$ eigenvalues. These first K eigenvalues correspond to the signal subspace while the remaining ones correspond to the noise subspace. Therefore, the covariance matrix can be decomposed as:

$$\mathbf{R}_\mathbf{y} = \mathbf{U}_\mathbf{x}\mathbf{\Lambda}_\mathbf{x}\mathbf{U}_\mathbf{x}^H + \mathbf{U}_\mathbf{n}\mathbf{\Lambda}_\mathbf{n}\mathbf{U}_\mathbf{n}^H \tag{2.30}$$

where $\mathbf{\Lambda}_\mathbf{x}$ is the $K \times K$ diagonal matrix containing the eigenvalues associated with the signal subspace and $\mathbf{U}_\mathbf{x}$ is the $M \times K$ matrix that contains the corresponding eigenvectors. Similarly, the eigenvalues of the noise subspace are contained in the matrix $\mathbf{\Lambda}_\mathbf{n}$ of size $(M-K) \times (M-K)$ and the corresponding eigenvectors in $\mathbf{U}_\mathbf{n}$. Since the noise subspace is orthogonal to the steering vectors corresponding to the true DOAs $\boldsymbol{\theta}$, it follows:

$$\mathbf{A}^H(\boldsymbol{\theta})\mathbf{U}_\mathbf{n} = \mathbf{0}_{K \times M-K}. \tag{2.31}$$

Equation (2.31) provides the MUSIC criterion for DOA estimation. In other words, the K minima of the projection $\mathbf{a}^H(\theta)\mathbf{U_n}\mathbf{U_n}^H\mathbf{a}(\theta)$ over the whole range of values for the parameter θ will correspond to the unknown DOAs. The MUSIC spectral estimate is then defined as:

$$P_{MUS}(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{U_n}\mathbf{U_n}^H\mathbf{a}(\theta)} \tag{2.32}$$

for which the unknown DOAs will now correspond to its K peaks. As in the case of conventional and MVDR beamformers, $\mathbf{R_y}$ is not available in practice and the sample covariance matrix is used instead. Its decomposition is given by:

$$\hat{\mathbf{R}}_\mathbf{y} = \hat{\mathbf{U}}_\mathbf{x}\hat{\mathbf{\Lambda}}_\mathbf{x}\hat{\mathbf{U}}_\mathbf{x}^H + \hat{\mathbf{U}}_\mathbf{n}\hat{\mathbf{\Lambda}}_\mathbf{n}\hat{\mathbf{U}}_\mathbf{n}^H \tag{2.33}$$

and hence, $\mathbf{U_n}$ should be replaced by its estimate $\hat{\mathbf{U}}_\mathbf{n}$ in (2.32).

As discussed in [13], $P_{MUS}(\theta)$ should be viewed more as a distance between two subspaces and not as a true spectrum estimate. Nevertheless, the MUSIC "pseudo-spectrum" still exhibits sharp peaks at the values of θ corresponding to the true DOAs. Similar to the beamforming methods, the algorithm requires a one dimensional search. However, there is an additional computational cost associated with the eigenvalue decomposition of the data covariance matrix. MUSIC provides a significant improvement in terms of estimation accuracy over the beamforming methods and when the time samples captured by each sensor of the array are sufficiently long, the algorithm provides statistically consistent estimates.

However, the main limitation of MUSIC appears in the so-called coherent source scenario. When some of the incoming signals happen to be highly correlated, then the algorithm's performance degrades dramatically. This is to be expected, as in that case the eigenvalue decomposition tends to underestimate the number of sources resulting in signal subspace estimate of reduced dimension. Similar issues can arise when the number of time snapshots are not sufficient enough. This problem is usually referred to as the rank-deficient case [26, 13].

2.1.4 Sparsity based approaches to DOA estimation

As stated previously, the MUSIC algorithm can achieve high resolution by focusing on a small number of "search directions" where the signals are present given the pseudo-

spectrum estimate. This implies some underlying sparsity in the spatial domain, considering that the algorithm is based on the assumption of a low dimensional signal subspace.

The emerging field of sparse representations has given renewed interest to the problem of source localization. The concept of spatial sparsity for DOA estimation was first introduced in [3], where it was shown that the source localization problem can be cast as a linear inverse problem, resulting in the development of the ℓ_1-SVD method. The specific method, assuming that only a few sources are present in contrast to the possible spatial locations, composes an overcomplete dictionary of steering vectors corresponding to each potential location of a source and estimates the unknown angles via ℓ_1 regularization. More recently, in a similar manner, this spatial sparsity property was also linked to the theoretical results of the sparse representations framework, utilizing a spatial compressed sensing approach for DOA estimation [27]. Interestingly, there is strong evidence that these sparsity based approaches can overcome the limitations of MUSIC algorithm while providing high resolution estimates. More specifically, the work in [3] showed that ℓ_1-SVD can achieve super-resolution even in the rank-deficient case (i.e. limitation of the data and/or correlation of the sources) and provide improved robustness to noise.

Since the aim of this thesis is to further investigate the problem of DOA estimation under spatial sparsity constraints and contribute to improving the specific approach, its detailed description will be put back until Chapter 5, where the problem is treated separately. However, the remaining sections of the current chapter serve as an introduction to the problems of sparse representation and compressed sensing, while at the same time they provide a detailed description of the main mathematical background on which this work is based.

2.2 Sparse representations

The concept of parsimony plays a principal role in many scientific areas of engineering and applied sciences. It implies that the description of a phenomenon by a simple model based on parsimonious terms is generally preferred over more complicated ones, since it can be more insightful. The field of signal processing is no exception to this principle. For the analysis of a family of signals or images, researchers more often than not intuitively employ transformations (e.g. orthonormal transforms such as the Fourier basis) from the

signal domain to some other domain, where the signals can be described or represented in a sparse and efficient way.

Sparse representations are signal expansions that can accurately represent the signal of interest using a linear combination of a relative small number of significant coefficients drawn from a basis or a redundant dictionary [28, 29]. In other words, the problem of finding sparse signal expansions can be redefined as a dimensionality reduction technique that looks for a compact expression involving only a few elementary components, which can reveal certain structures of the given signal.

The problem of sparse representations lies at the core of almost every aspect of signal processing, with a broad range of applications including denoising [30], deblurring [31], compression [32] and many more. For instance, in compression, the size of a signal can be reduced by coding only a few nonzero coefficients of the sparse representation [28]. Thus, the sparser a representation the higher the signal-compression factor achieved, resulting in reduction of the storage requirements or in faster transmission. It follows that identifying the appropriate transform basis that can yield sparser representations or better approximations is of vital importance.

Initially, researchers were mainly focused on the design of orthonormal basis dictionaries due to their attractive properties. As nicely put in [33], orthonormal transformations can be viewed as rotations in a high dimensional space, and hence their inversion is trivial (e.g. by applying the adjoint operator). Consequently, the ease with which the data can be transformed from the signal to the transform domain and vice versa makes orthonormal bases very convenient and effortless to apply in practice. This is the main reason why transforms such as the Discrete Cosine Transform (DCT), which is a variant to the discrete Fourier transform (DFT) and Discrete Wavelet Transform (DWT) have been used extensively over the past decades and eventually found their place in industry standards for compression, such as MPEG-1, MPEG-2 and JPEG-2000.

However, not all signals can be efficiently represented as a linear combination of elements drawn from an orthonormal basis. In fact, there exist complex signals, which can be mixtures of different phenomena and for this reason they require different transforms to be accurately described [34]. This has led many researchers to form redundant dictionaries, typically consisting of unions of orthogonal bases (e.g. concatenation of

FFT and Dirac basis), which can more efficiently capture the information present in the signal of interest and thus allow for sparser representations or better approximations [28, 35]. Furthermore, this redundancy suggests better adaptability and more flexibility in terms of the sparse representation or the signal approximation. Mallat in [28] very aptly draws the parallel between these redundant or overcomplete dictionaries, which are a collection of simple waveforms and language dictionaries. As a richer dictionary helps people express their ideas in shorter phrases, it has become widely accepted that a more redundant dictionary can lead to better and more compact signal descriptions. For instance, FFT may not be sufficient for the analysis of time varying signals such as music or speech. Redundant transforms (e.g. Gabor transform or STFT) play an important role for the analysis of these types of signals, as they lead to localized time-frequency representations. Since audio signals can be viewed as the variation of frequency events in time, time-frequency representations can better capture their properties. Considering also that overcomplete time-frequency transforms can result in sparse representations of audio signals, the appropriate thresholding of such representations of noisy audio signals could yield efficient denoising algorithms. In fact, redundancy has been proven to be very beneficial in applications, such as denoising or deblurring. It has also paved the way for sparse representations to find applicability in many other signal processing applications, such as feature extraction [36, 37] and source separation [38, 39] opening new horizons for scientific research.

Another emerging topic in the field of sparse representations that has gained much interest recently is dictionary learning (DL). DL is concerned with learning the redundant dictionary from the data, instead of building the dictionary as a concatenation of several bases. In that way, the dictionary is designed to better fit the model by adapting the dictionary to a set of training signals [40].

However, this flexibility of choosing elements from the dictionary comes with a price to pay. In contrast with the straightforward orthonormal transformation case, finding the sparsest representation in a redundant dictionary is a non-trivial combinatorial problem that cannot be solved in polynomial time (i.e. NP-hard) [41], as it involves the exhaustive search over all possible combinations of dictionary vectors. Nevertheless, as will be discussed in following sections, suboptimal methods have been developed, for which

rigorous theoretical results exist that can guarantee optimality under certain conditions.

2.2.1 The pursuit of a sparse representation

Let $\mathbf{y} \in \mathbb{R}^M$ be an observed vector to be decomposed and represented in the dictionary $\mathbf{A}$ using a small number K of significant coefficients corresponding to the columns of the full rank matrix $\mathbf{A}$. If the dictionary is a square and nonsingular matrix (e.g. Fourier or wavelets basis) and a sparse representation of $\mathbf{y}$ exists in the specific basis, the sparse representation can be easily estimated by a single matrix inversion. However, as explained previously redundant or overcomplete dictionaries will result in sparser decompositions. In that case, $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M < N$ and subsequently the number of columns is larger than the dimensionality of the observed signal's space. Therefore, in the noiseless case, given $\mathbf{y}$ and $\mathbf{A}$, the following ill-posed inverse problem for the unknown vector $\mathbf{x}$ has to be solved:

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.34}$$

where $\mathbf{x} = [x_1, \ldots, x_N]^T$ is K-sparse, namely it has at most $K \geq \|x\|_0$ nonzero entries, with $K \ll N$. $\|\cdot\|_0$ denotes the ℓ_0 norm defined as:

$$\|\mathbf{x}\|_0 = |\text{supp}(\mathbf{x})| \quad \text{where} \quad \text{supp}(\mathbf{x}) = \{1 \leq i \leq N : x_i \neq 0\} \tag{2.35}$$

and $|\cdot|$ denotes the cardinality, namely the number of elements of a set.

The above system of linear equations (2.34) is said to be an underdetermined system, as the number of unknowns or variables is larger than the number of equations. Such a system yields an infinite number of solutions. In sparse coding, the main interest is to obtain the sparsest solution, which is the one that has the smallest number of nonzero elements [42]. This implies minimization of the ℓ_0 norm and subsequently the sparse representation problem can be formulated as:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{2.36}$$

The problem of (2.36) is concerned with the noiseless case. In the noisy case, the observed vector $\mathbf{y}$ is corrupted by additive noise such that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$. Accounting for the noise,

the optimization problem of (2.36) becomes:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{such that} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \tag{2.37}$$

where ϵ is a small constant depending on the noise level.

2.2.2 Uniqueness of the sparse recovery problem

It has been shown so far that the sparse recovery problem in redundant dictionaries is not as straightforward as dealing with orthonormal bases and additional constraints need to be employed. However, this does not ensure that the solution obtained by solving problem (2.36) will be the sparsest possible representation. Donoho et al. in [43] provide the necessary and sufficient condition for uniqueness of the sparsest representation.

The spark of a matrix $\mathbf{A}$ is defined as the smallest possible number of linear dependent columns from $\mathbf{A}$:

$$\text{spark}(\mathbf{A}) = \min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{such that} \quad \mathbf{A}\mathbf{z} = \mathbf{0}. \tag{2.38}$$

Although the spark is in a way related to the rank of a matrix $\mathbf{A}$, it is quite different, since the rank is defined as the maximal number of columns from $\mathbf{A}$ that are linear independent. However, it holds that $\text{spark}(\mathbf{A}) \leq \text{rank}(\mathbf{A}) + 1$. Using this definition of the spark, the authors in [43], proved that an upper bound exists on the sparsity level K, so that the solution to the sparse recovery problem of (2.36) is unique.

Theorem 1. (Donoho & Elad [43]) *A representation* $\mathbf{x}$ *of* $\mathbf{y}$ *is the sparsest possible if*

$$\|\mathbf{x}\|_0 < \frac{\text{spark}(\mathbf{A})}{2}. \tag{2.39}$$

Proof. Assume that there exist two sparse representations $\mathbf{x}_1$ and $\mathbf{x}_2$ of $\mathbf{y}$ with $\mathbf{x}_1 \neq \mathbf{x}_2$, both satisfying (2.39). It follows:

$$\begin{aligned} &\mathbf{y} = \mathbf{A}\mathbf{x}_1 = \mathbf{A}\mathbf{x}_2 \Rightarrow \\ &\mathbf{A}\mathbf{x}_1 - \mathbf{A}\mathbf{x}_2 = \mathbf{0} \Rightarrow \\ &\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0} \Rightarrow \\ &\|\mathbf{x}_1 - \mathbf{x}_2\|_0 < \text{spark}(\mathbf{A}) \end{aligned} \tag{2.40}$$

which is contradictory according to the definition of spark. Therefore, it can only be that $\mathbf{x}_1 = \mathbf{x}_2$. □

2.2.3 Algorithms for sparse recovery

As long as condition (2.39) of Theorem (1) holds, it is guaranteed that the sparse recovery problem of (2.34) has a unique solution. However, the optimization in (2.36) is not convex and therefore its solution is computationally intractable, since it is known to be NP-hard problem [41].

Instead of solving the ℓ_0 minimization problem, other suboptimal strategies are adopted in practice. According to the method of ℓ_1 minimization or Basis Pursuit (BP), [44] one can attempt to solve the convex optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{2.41}$$

which when compared to (2.36), simply replaces the ℓ_0 norm with the ℓ_1 norm, defined as the sum of the absolute values of the sparse vector $\mathbf{x}$. The problem of ℓ_1 norm minimization is convex and can be solved using linear programming (LP) methods e.g. interior-point methods in polynomial time [45]. As explained in [46], in contrast to ℓ_2 norm regularization, minimization of the ℓ_1 norm promotes sparsity and therefore yields sparse solutions. This is illustrated geometrically in Fig. 2.3, which shows plots of ℓ_p balls for different values of p ($p = 1$, $p = 2$, $p = \infty$ and $p = 1/2$) in $\mathbb{R}^2$. The red hyperplane corresponds to the nullspace of the one dimensional subspace $\mathbf{A}$. As can be seen, only the ℓ_p balls with $p \leq 1$ (Fig. 2.3(a) and Fig. 2.3(d)) promote sparse solutions. On the other hand, ℓ_p balls with $p > 1$ (Fig. 2.3(b) and Fig. 2.3(c)) spread the solution to both coefficients.

Nevertheless, although BP is computationally tractable, its convergence is rather slow and other faster greedy algorithms such as Matching Pursuit (MP) [47] and Orthogonal Matching Pursuit (OMP) [48] can serve as alternatives. Both MP and OMP algorithms are iterative in nature, selecting one atom from the dictionary $\mathbf{A}$ at a time; the one that is most correlated to the residual, which is obtained by subtracting the contribution of a partial estimate $\hat{\mathbf{y}}$ of the measurements from $\mathbf{y}$. The correlations are computed by projecting the current residual onto the columns of the dictionary $\mathbf{A}$. OMP provides an

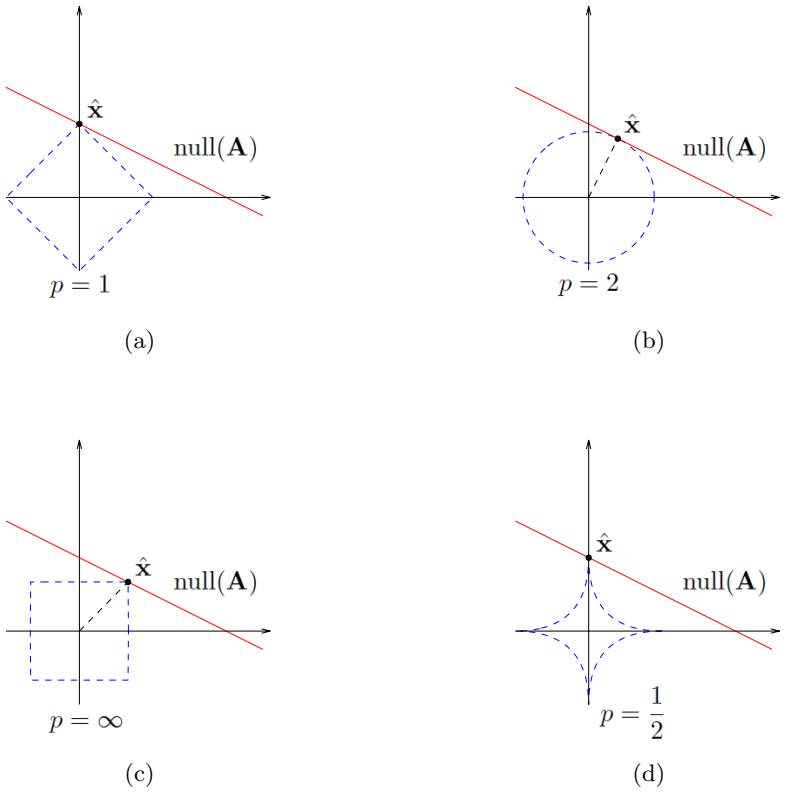


Figure 2.3: Geometry of the ℓ_p recovery: Best approximation of a point in $\mathbb{R}^2$ by a one dimensional subspace using the ℓ_p balls: (a) ℓ_1 ball, (b) ℓ_2 ball, (c) ℓ_∞ ball and (d) quasinorm ball $p = 1/2$.

improved version to MP by orthogonalizing the directions of projection at each iteration. The main steps of the OMP algorithm are summarized in Algorithm 1.

Similarly, in the noisy case, for Basis Pursuit Denoise (BPDN) the ℓ_0 norm in (2.37) is replaced by the ℓ_1 norm and the resulting optimization problem is solved. For OMP the stopping criterion should be adjusted accordingly [49] to compensate for the additive noise in the measurements.

BP and OMP stand as the most popular sparse recovery algorithms. However, it is generally believed that in certain cases BP provides superior recovery performance [44, 50]. For this reason, many other approaches have been proposed in literature including Lasso [51], LARS [52] and GPSR [53] that attempt to bridge the gap between computational simplicity and optimality. Most of these algorithms attempt to replace the local optimization performed by OMP with a global optimization criterion and at the same time retain the fast iterative nature of greedy algorithms.

Algorithm 1 Orthogonal Matching Pursuit

Input: $\mathbf{A}$, $\mathbf{y}$
Set stopping conditions $l_{\max}$ and $\epsilon_{\min}$
Initialize: $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{x}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\|\mathbf{r}^k\|_2 > \epsilon_{\min}$ **do** {Find next atom}
    $k \leftarrow k + 1$
    $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} |\mathbf{a}_i^T \mathbf{r}^{k-1}|$
    Add constraints:
    $\mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\},\ \mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{a}_{i^k}]$
    $\mathbf{x}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{y},\ \hat{\mathbf{y}}^k \leftarrow \mathbf{A}^k \mathbf{x}^k,\ \mathbf{r}^k \leftarrow \mathbf{y} - \hat{\mathbf{y}}^k$
end while
Output: $\mathbf{x}^* = \mathbf{x}^k$

2.2.4 Recovery guarantees

Both ℓ_1 minimization and OMP are suboptimal methods to the combinatorial problem of ℓ_0 minimization. The question that naturally arises is how well these algorithms can approximate the sparsest solution or under which conditions the BP and OMP solutions will be equivalent to the solution of (2.36). This topic has received much attention within the community of sparse representations and there exist theoretical results that can guarantee optimality for both algorithms.

The problem of ℓ_1/ℓ_0 equivalence was initially studied by Donoho and Huo [54]. In the specific work, it was shown that under some stronger condition than the ℓ_0 uniqueness condition of (2.39), ℓ_1 minimization finds the optimal sparsest solution. Their result is summarized in the following theorem:

Theorem 2. (Donoho & Huo [54]) *Let* $\mathbf{A}$ *be a union of two orthonormal bases with mutual coherence* μ*. If for the sparse representation* $\mathbf{x}$ *of* $\mathbf{y}$ *holds that*

$$\|\mathbf{x}\|_0 < \frac{1}{2}(\mu^{-1} + 1) \tag{2.42}$$

then $\mathbf{x}$ *is the unique solution to both* ℓ_1 *and* ℓ_0 *minimization problems.*

In Theorem 2, the dictionary mutual coherence refers to the maximum correlation between any two dictionary atoms. It is expressed as the largest absolute value of the inner product between any two different and normalized column vectors of the matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$:

$$\mu = \max_{1 \le i, i \ne j \le N} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\| \|\mathbf{a}_j\|} \tag{2.43}$$

where $\mathbf{a}_i$ denotes the i-th column of the dictionary. Hence, the coherence is upper bounded by the value 1 ($\mu \leq 1$), which corresponds to a coherent dictionary and lower bounded by the inequality $\mu \geq \sqrt{\frac{N-M}{M(N-1)}}$. A dictionary that meets this lower bound is called an optimal Grassmannian frame [55]. According to the condition (2.42) of Theorem 2 the lower the mutual coherence of the dictionary, the larger the bound on the sparsity level K of $\mathbf{x}$. This is something that should be expected, since low coherence implies that the dictionary is close to orthogonal and therefore the K subspaces defined by a subset of its columns can be easily distinguished from other subspaces of the same dimensionality resulting in efficient recovery of the K-sparse vector $\mathbf{x}$. On the other hand, if $\mathbf{A}$ contains highly correlated atoms and subsequently its coherence μ is high (close to 1), then there is no guarantee that ℓ_1 minimization will efficiently recover the solution vector. This is the reason why in sparse representations incoherent dictionaries are always embraced.

Theorem 2 is restricted to redundant dictionaries formed from union of bases. However, additional work on the topic of ℓ_1/ℓ_0 equivalence [56, 57, 43, 58] has shown that this result can be generalized for any redundant dictionary. Regarding OMP, similar theoretical guarantees exist. More specifically, Tropp in [35] showed that the following theorem holds.

Theorem 3. (Tropp [35]) *Let* $\mathbf{y}$ *be the signal to be decomposed in the redundant dictionary* $\mathbf{A}$ *and* S *be the support of the sparsest solution vector* $\mathbf{x}$ *(*$S = supp(\mathbf{x})$*). If for any column vector* $\mathbf{a}_j$ *with* $j \notin S$ *holds*

$$\max_{j \notin S} \|\mathbf{A}_S^{\dagger}\mathbf{a}_j\|_1 < 1 \tag{2.44}$$

OMP will recover the sparsest vector $\mathbf{x}$*.*

The result of Theorem 3 is the best possible for OMP algorithm and the condition (2.44) is referred to as Exact Recovery Condition (ERC). As also shown in [35], ERC implies the following theorem.

Theorem 4. (Tropp [35]) *ERC holds for every sparse representation* $\mathbf{x}$ *whenever* $\|\mathbf{x}\|_0 < \frac{1}{2}(\mu^{-1} + 1)$.

Theorem 4 reveals that condition (2.42) guarantees optimality for both BP and OMP.

It can be easily shown that the specific condition is stronger than the ℓ_0 uniqueness condition of (2.39), since the lower bound on the spark of dictionary is spark($\mathbf{A}$) $\geq 1/\mu+1$ [56, 43, 46].

The above theoretical result is identical for BP and OMP algorithms and it largely agrees with empirical results showing that both algorithms offer comparable performance. However, there have been reported theoretical results, which demonstrate that in certain scenarios BP succeeds in finding the optimal solution, whereas OMP fails. As also discussed in [35], ERC provides the best possible condition for OMP but not always for BP. There are cases, for which ℓ_1 minimization recovers the sparsest solution even if ERC is violated [28]. The exact recovery for ℓ_1 minimization can be replaced by a more precise sufficient criterion introduced by Fuchs in [59]. The so-called Fuchs Condition will be discussed in more detail in Chapter 3.

Recovery conditions based on the mutual coherence of the dictionary for the noisy case also exist [60] for both BP and OMP algorithms, that are mostly in favour of ℓ_1 minimization.

2.3 Compressed sensing

The extensive study of the sparse representation problem and more general underdetermined inverse problems, apart from bringing new theoretical insights regarding the optimality of suboptimal regularization methods, has also opened new avenues for scientific research in several signal processing applications. Among those, compressed sensing or compressive sampling (CS) has attracted significant attention within the signal processing community.

Sampling is a subfield of digital signal processing and it is concerned with the conversion of a continuous time signal into a discrete sequence of numbers. The most widely known principle in sampling and data acquisition is the Shannon-Nyquist theorem [61], which states that if a band-limited signal $x(t)$ has no frequency components higher than f_x Hz, it can be perfectly reconstructed from a set of samples taken at the rate of f_s ($\geq 2f_x$) samples per second. CS is a new technique, which goes against this principle and does not require that a signal must be sampled at a rate at least twice its highest frequency. According to this method, initially introduced by Donoho [62] and Candès

[63], one can sample at a greatly lower rate than traditional methods suggest and under certain conditions the optimal sparse representation that allows perfect reconstruction of the original signal can be estimated [64].

The underlying assumption of CS is that the signal is sparse or compressible in some domain and subsequently it can be accurately or approximately expressed as a linear combination of a small number of basis functions [65]. Note that unlike the previous discussion, sparsity here is assumed over an orthogonal transform. This sparsity or compressibility of the signal plays a key role in the CS concept, as it leads to dimensionality reduction and efficient modeling. In the classical compression approach, the signal is first sampled densely and then transformed in a domain that is sparse. This allows compression of the signal by encoding a small number of transform coefficients, which retain most of the information existing in the signal, while the remaining coefficients are discarded.

CS asserts that this two stage approach of sampling and compression can be merged into one. Thus, given some highly undersampled data that capture most or all of the useful information, one can design robust and efficient data acquisition protocols at a much lower computational cost [63]. These undersampled data correspond to a reduced set of linear and non-adaptive measurements [66, 67]. Among the practical benefits is that the measurement or sensing matrix can be a random projection. According to the CS theory and under certain conditions, this sensing mechanism except for reducing the dimensionality, preserves all or most of the information when the signal is said to be sparse or compressible, respectively.

2.3.1 The compressed sensing recovery problem

Consider a linear measurement sensor that obtains M samples or observations of the discrete signal of interest $\mathbf{f} \in \mathbb{R}^N$ with $M < N$. The observations or measurements vector can be written as:

$$\mathbf{y} = \mathbf{\Phi f} \tag{2.45}$$

where $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ represents the measurement process.

The main assumption is that the discrete signal $\mathbf{f}$ is sparse in some domain and can be decomposed and accurately represented in the basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ using the following

equation:

$$\mathbf{f} = \mathbf{\Psi x} \tag{2.46}$$

where $\mathbf{x}$ is a K-sparse representation, meaning that it has only a small number of K nonzero entries, with $K \ll N$.

The CS recovery problem can be formulated by combining equations (2.45) and (2.46). The resulting system of linear equations is:

$$\mathbf{y} = \mathbf{\Phi\Psi x} = \mathbf{Ax} \tag{2.47}$$

where the redundant dictionary $M \times N$ dictionary $\mathbf{A}$ is given by $\mathbf{A} = \mathbf{\Phi\Psi}$. Therefore, as in the case of finding a sparse representation, CS requires the inversion of an underdetermined system of linear equations with the additional constraint of sparsity. It follows that for the CS framework the two critical points are sparsity and incoherence. For a given sparsity level K, CS is mainly concerned with low coherence pairs, namely matrices that contain low correlated elements. If the measurement matrix $\mathbf{\Phi}$ and the representation matrix $\mathbf{\Psi}$ are sufficiently incoherent and the number of measurements is sufficiently large:

$$M \geq C\mu^2 K \log N \tag{2.48}$$

for some constant C, the sparse vector $\mathbf{x}$ and consequently the original signal $\mathbf{f}$ can be exactly recovered with high probability by ℓ_1 minimization. The coherence μ of the overcomplete matrix $\mathbf{A}$ can be viewed as a measure of incoherence between the sensing matrix $\mathbf{\Phi}$ and the representation matrix $\mathbf{\Psi}$. As can be seen from equation (2.48), a lower coherence μ (or equivalently higher incoherence) leads to a smaller lower bound on the number of measurements M for a fixed sparsity level K. Therefore, an incoherent pair of matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ will require fewer measurements for the exact recovery of $\mathbf{f}$. Random matrices is known to be largely incoherent with any fixed basis that can lead to a sparse representation. Although the random projection matrix $\mathbf{\Phi}$ is not square, it preserves the structure and information present in the signal of interest, when the signal is sparse [68].

When dealing with not exactly sparse signals, it becomes clear that one cannot obtain perfect recovery of the signal $\mathbf{f}$. In that case, CS should be viewed as a lossy compression scheme. This can arise in several real world problems, since real world signals are not truly

exact sparse. However, CS can still be applied if these signals are at least compressible in some transform domain, meaning that the nonzero entries of the unknown vector $\mathbf{x}$ follow some fast exponential decay [69] and therefore there are only a few significant coefficients. Although information loss is inevitable, a good approximation of $\mathbf{f}$ can still be obtained.

2.3.2 The Restricted Isometry Property

The CS recovery problem can be viewed as a special case to the more general sparse recovery problem. Nevertheless, the advent of CS has brought to the world of sparse representations the concept of random overcomplete matrices, which in turn has resulted in further theoretical advances that have been proven to be beneficial for both CS and sparse representations.

Among those, the Restricted Isometry Property (RIP) [70] from a theoretical perspective has been proven to be a very powerful tool that has played an important role in popularizing the CS framework. Given a redundant dictionary $\mathbf{A}$, RIP can be used to answer the question of how near to orthogonal is the column space of $\mathbf{A}$. More specifically, according to the RIP, if a matrix $\mathbf{A}$ obeys:

$$(1-\delta_K)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\delta_K)\|\mathbf{x}\|_2^2 \tag{2.49}$$

where δ_K is a small constant often referred to as the restricted isometry constant (RIC) and if $\delta_{2K} < \sqrt{2}-1$ then the K-sparse vector $\mathbf{x}$ cannot be in the nullspace of $\mathbf{A}$. In other words, if the RIP holds for a matrix $\mathbf{A}$ then all subset of K columns taken from $\mathbf{A}$ will be nearly orthogonal. Therefore, RIP characterizes matrices that are nearly orthogonal when operating on sparse vectors. It also implies that these matrices will exhibit low coherence. To see the connection between coherence and RIP, consider the normalized dictionary $\mathbf{A}$. If its mutual coherence is μ, then $\mathbf{A}$ satisfies the RIP of order K with $\delta = (K-1)\mu$ for all $K < 1/\mu$ [67].

RIP is a sufficient condition that when satisfied by the dictionary $\mathbf{A}$, it guarantees the full recovery of a sparse vector. When the RIP holds given a CS recovery problem in the noiseless case and $\delta_{2K} < \sqrt{2}-1$, then the solution $\hat{\mathbf{x}}$ obtained by ℓ_1 minimization

obeys:

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq C_0 K^{-1/2} \|\mathbf{x} - \mathbf{x}_K\|_1 \tag{2.50}$$

for some constant C_0, where $\mathbf{x}_K$ is the vector $\mathbf{x}$ with all but the largest K components set to zero. Accordingly, in the noisy case for BPDN, it holds:

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq C_0 K^{-1/2} \|\mathbf{x} - \mathbf{x}_K\|_1 + C_1 \epsilon \tag{2.51}$$

where C_0 and C_1 are constants and ϵ is the constant that bounds the noise and depends on the noise level.

It has been proven that for matrices containing random or partial Fourier entries RIP holds with high probability [71]. More specifically, when the redundant dictionary $\mathbf{A}$ is randomly generated by sampling its columns from the uniform, normal or Bernoulli distributions, then it will obey the RIP as long as:

$$M \geq CK \log(N/K) \tag{2.52}$$

where C is a constant depending on each instance [68]. The case of a partial random Fourier ensemble has been also extensively studied and similar results have been shown to exist [72]. In particular, for a partial Fourier dictionary $\mathbf{A}$ the RIP is satisfied with large probability provided that:

$$M \geq CK(\log N)^4. \tag{2.53}$$

The result of equation (2.53) for partial Fourier dictionaries is of great importance to the problem of DOA estimation. As will be discussed in more detail in Chapter 5, under the appropriate discretization of the angular space and spacing of the array sensors the DOA dictionary draws resemblance to a partial Fourier matrix.

Note that the lower bounds of (2.52) and (2.53) obtained using an RIP based theoretical analysis slightly differ from the result of (2.48) that relies solely on the mutual coherence of the dictionary. Furthermore, for a matrix that obeys the RIP, the approximation error is given by the inequality in (2.50) or (2.51). Despite the fact that similar coherence based guarantees for ℓ_1 minimization also exist [60], the advantage of the re-

sults based on RIP is because they are deterministic and universal. The same does not hold for the coherence based guarantees. As can be seen, the lower bounds on the number of measurements M (2.52) and (2.53) are free from any stochastic quantities. However, the inequality of equation (2.48) depends on the mutual coherence which is a stochastic bound in nature, since different dictionaries will exhibit different maximal incoherence.

2.4 Discussion

In this chapter, the problem of narrowband DOA estimation in sensor arrays was introduced. After describing the signal model and the resulting parameter estimation problem, the non-parametric array processing techniques for DOA estimation were presented. The main idea and motivation behind the formulation of the DOA estimation problem as an inverse sparse recovery problem was also briefly discussed. This discussion will continue in following chapters, since it constitutes the main focus of this thesis.

Next, the basic concepts of the sparse representation problem have been presented. After highlighting the main advantages of using redundant transforms over conventional orthonormal bases for signal representation, this chapter provided a concise overview of the most popular algorithms for solving the sparse recovery problem along with vigorous theoretical guarantees that justify their recovery capabilities. The compressed sensing framework was also briefly described. The theoretical analysis based on the RIP not only verifies the robustness of CS as a compression and data acquisition protocol, but it also shows its potential to provide further insight to the more general problem of sparse recovery.

Chapter 3

Sparse representations & compressed sensing recovery algorithms for large scale problems

---

In the previous chapter, the main concept of sparse representations was introduced and briefly described. Before investigating the applicability of the sparse representation framework to the problem of DOA estimation, a more in depth look at the problem of sparse recovery is given, which in turn leads to the development of a novel suite of algorithms that tackle this problem.

As mentioned in Chapter 2, BP possesses certain theoretical properties that justify its superior performance over greedy algorithms such as OMP. Nevertheless, its convergence is rather slow and for this reason in practical signal processing applications OMP has been used extensively due to its fast convergence, which comes as a result of its algorithmic simplicity. This observation has driven the motivation to develop algorithms with good theoretical properties, as in the case of ℓ_1 minimization, and with the ability to achieve faster convergence exhibiting reduced complexity, as happens in the case of OMP. Polytope Faces Pursuit (PFP) is a greedy approach that meets these requirements. The specific algorithm, initially introduced by Plumbley in [1], solves the sparse recovery problem performing BP with similar order complexity to OMP. PFP adds one vector basis at a time and adopts a path following approach based on the geometry of the polar polytope associated with the dual linear program.

In this chapter, the PFP algorithm is first introduced and then an alternative ap-

proach to its original version of updating the solution vector at each step using directional updates is proposed. As in the case of OMP, Cholesky factorization, adopted by the original version of PFP, requires storage of large matrices which can be expensive when it comes to large scale recovery problems. More specifically, the proposed algorithm of Conjugate Gradient Polytope Faces Pursuit uses the very well known method of conjugate gradients to update the solution vector thereby reducing the storage requirements of the original algorithm. Theoretical and empirical results demonstrate the robustness of the proposed sparse recovery method and outline its advantages and shortcomings.

Then, alternative stagewise dictionary atom selection strategies that can further reduce the computational complexity are also considered. The resulting Stagewise Conjugate Gradient Polytope Faces Pursuit algorithm adds several column vectors at each stage according to a fixed or weak atom selection criterion.

3.1 Overview of Polytope Faces Pursuit

The traditional ℓ_1-minimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{3.1}$$

can be converted to its standard form using nonnegative coefficients [2]:

$$\min_{\tilde{\mathbf{x}}} \mathbf{1}^T\tilde{\mathbf{x}} \quad \text{such that} \quad \mathbf{y} = \tilde{\mathbf{A}}\tilde{\mathbf{x}},\ \tilde{\mathbf{x}} \geq 0 \tag{3.2}$$

where $\mathbf{1}$ is a column vector of ones, $\tilde{\mathbf{A}} = [\mathbf{A}, -\mathbf{A}]$ and $\tilde{\mathbf{x}}$ is the $2N$ nonnegative vector:

$$\tilde{x_i} = \begin{cases} \max(x_i, 0) & 1 \leq i \leq N \\ \max(-x_{i-N}, 0) & N+1 \leq i \leq 2N \end{cases}. \tag{3.3}$$

The primal linear program (3.2) has a corresponding dual linear program:

$$\max_{\mathbf{c}} \mathbf{y}^T\mathbf{c} \quad \text{such that} \quad \tilde{\mathbf{A}}^T\mathbf{c} \leq \mathbf{1} \tag{3.4}$$

such that a bounded solution to (3.4) exists if and only if a bounded solution to (3.2) exists. Thus, one can initially look for a solution $\mathbf{c}^*$ to (3.4) and use the Karush-Kuhn-

Tucker (KKT) [2] conditions to solve the resulting system for $\tilde{\mathbf{x}}^*$. More formally, the following lemma provides the optimality conditions for this system.

Lemma 1. (Plumbley [73]) *Suppose that the primal problem is feasible. Then the pair* $\tilde{\mathbf{x}}^*$, $\mathbf{c}^*$ *is an optimum point for both primal and dual linear programs if and only if the following conditions hold:*

$$\tilde{\mathbf{A}}\tilde{\mathbf{x}}^* = \mathbf{y},\ \tilde{\mathbf{x}}^* \geq \mathbf{0} \tag{3.5a}$$

$$\tilde{\mathbf{A}}^T\mathbf{c}^* \leq \mathbf{1} \tag{3.5b}$$

$$(\tilde{\mathbf{a}}_j^T\mathbf{c}^* - 1)\tilde{x}_j^* = 0 \quad \forall j, 1 \leq j \leq 2N. \tag{3.5c}$$

The work in [74] looks at the dual linear program (3.4) from a geometrical perspective, using the geometry of polytopes. The set $P^* = \{\mathbf{c} \mid \pm\mathbf{a}_i^T\mathbf{c} \leq 1, \mathbf{a}_i \in \mathbf{A}\}$ of feasible solutions to the inequality $\tilde{\mathbf{A}}^T\mathbf{c} \leq \mathbf{1}$ defines the so-called polar polytope (bounded M-dimensional polygon/polyhedron) corresponding to the region bounded by that inequality. P^* is dual to the primal polytope $P = \text{conv}\{\pm\mathbf{a}_i, \mathbf{a}_i \in \mathbf{A}\}$, which is the convex hull corresponding to the doubled dictionary matrix, as illustrated in Fig. 3.1 (dashed polytope).

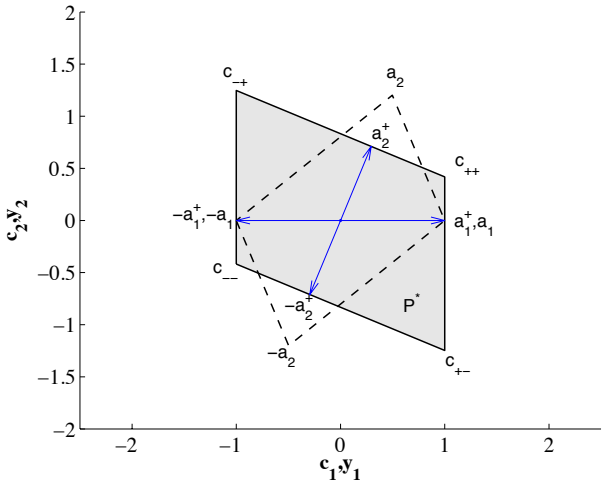


Figure 3.1: Example of a 2-D polytope [1]: primal (dashed) and polar (solid) polytopes.

The scaled vectors $\pm\mathbf{a}_i^+ = \mathbf{a}_i/\|\mathbf{a}_i\|_2$ touch the faces of P^* and the vertices $\mathbf{c}_{++}$, $\mathbf{c}_{-+}$, $\mathbf{c}_{+-}$, $\mathbf{c}_{--}$ correspond to particular sets of selected atoms (Fig. 3.1). It is a standard result from linear programming that the optimum $\mathbf{c}^*$ will be achieved at one of those

vertices [75]. Therefore, the Polytope Faces Pursuit algorithm operates in the polar polytope and attempts to find this optimum $\mathbf{c}^*$. The algorithm starts at $\mathbf{c} = 0$ and adopts a path following approach towards the measurement vector $\mathbf{y}$ until it hits a face of P^*. The next face encountered is the one along the current face towards the projected residual. Hence, for the selection of the next atom, PFP at the k-th iteration follows the path $\mathbf{h} = \mathbf{c}^{k-1} + \lambda \mathbf{r}^{k-1}$, where $\mathbf{r}^{k-1} = \mathbf{y} - \hat{\mathbf{y}}^{k-1}$ is the current residual with $\hat{\mathbf{y}}^{k-1}$ being the current estimate, and requires that $\tilde{\mathbf{a}}_{i^k}^T \mathbf{h} = 1$. Accordingly, the next face will

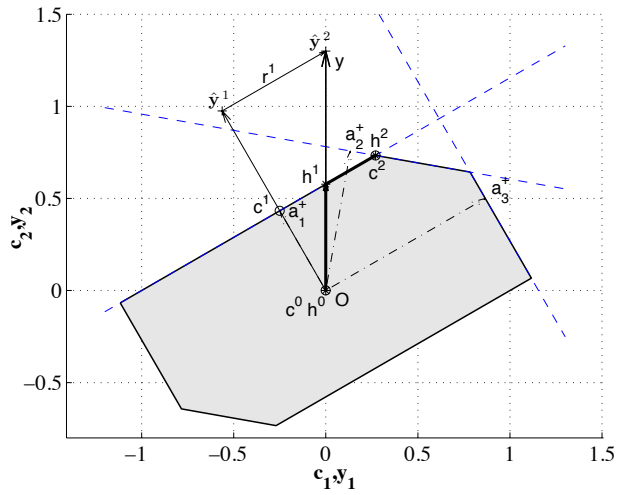


Figure 3.2: Path of the Polytope Faces Pursuit algorithm [1].

be encountered at the minimum positive value of λ such that:

$$\lambda_k = \min_{i \notin \mathcal{I}^{k-1}} \{\lambda > 0 \mid \tilde{\mathbf{a}}_i^T (\mathbf{c}^{k-1} + \lambda \mathbf{r}^{k-1}) = 1\} = \min_{i \notin \mathcal{I}^{k-1}} \{\frac{1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}}{\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1}} \mid \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > 0\}. \quad (3.6)$$

It becomes apparent that at each step PFP uses the maximum scaled correlation (defined as the inverse of λ_k):

$$\alpha_k = \arg \max_{i \notin \mathcal{I}^{k-1}} \frac{\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1}}{1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}} \quad (3.7)$$

as an atom selection criterion, where atoms $\tilde{\mathbf{a}}_i$ for which $\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > 0$ are only considered. Atoms that have already been selected are also excluded i.e. $i \notin \mathcal{I}^{k-1}$, where $\mathcal{I}^k$ denotes the set of indices corresponding to the already selected atoms at the k-th iteration. Once a new face has been identified, its index i^k is added to the current set i.e. $\mathcal{I}^k = \mathcal{I}^{k-1} \cup \{i^k\}$. After updating the solution vector the algorithm requires releasing of certain constraints, namely switching out of a basis vector $\tilde{\mathbf{a}}_j$ whenever its corresponding entry $\tilde{x}_j^k$ is negative

i.e. $\mathcal{I}^k = \mathcal{I}^k \setminus \{j\}$.

Fig. 3.2 shows a polar polytope and its optimum basis vertex $\mathbf{c}^*$, which as expected is the furthest vertex along the direction of $\mathbf{y}$. The full Polytope Faces Pursuit algorithm is given in [1] and illustrated in Algorithm 2.

Algorithm 2 Polytope Faces Pursuit (original version)

Input: $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_i]$, $\mathbf{y}$ {If required, set $\tilde{\mathbf{A}} \leftarrow [\mathbf{A}, -\mathbf{A}]$}
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize: $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \tilde{\mathbf{A}}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \tilde{\mathbf{x}}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next face}
  $k \leftarrow k + 1$
  Find face: $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \{(\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1})/(1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}) \mid \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > 0\}$
  Optionally: $\lambda_k \leftarrow (1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1})/(\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1})$
  Add constraints: $\tilde{\mathbf{A}}^k \leftarrow [\tilde{\mathbf{A}}^{k-1},\ \tilde{\mathbf{a}}_{i^k}],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\tilde{\mathbf{x}}^k \leftarrow \tilde{\mathbf{A}}^{k\dagger}\mathbf{y}$
  while $\tilde{\mathbf{x}}^k \not\geq \mathbf{0}$ **do** {Release retarding constraints}
    Select some $j \in \mathcal{I}^k$ such that $\tilde{x}_j^k < 0$; remove column $\tilde{\mathbf{a}}_j$ from $\tilde{\mathbf{A}}^k$
    Update: $\mathcal{I}^k \leftarrow \mathcal{I}^k \setminus \{j\},\ \tilde{\mathbf{x}}^k \leftarrow \tilde{\mathbf{A}}^{k\dagger}\mathbf{y}$
  end while
  $\mathbf{c}^k \leftarrow (\tilde{\mathbf{A}}^{k\dagger})^T \mathbf{1},\ \hat{\mathbf{y}}^k \leftarrow \tilde{\mathbf{A}}^k \tilde{\mathbf{x}}^k,\ \mathbf{r}^k \leftarrow \mathbf{y} - \hat{\mathbf{y}}^k$
end while
Output: $\mathbf{c}^* = \mathbf{c}^k$, $\tilde{\mathbf{x}}^* \leftarrow \mathbf{0}$ + corresponding entries from $\tilde{\mathbf{x}}^k$
{If required, get $x_i^* \leftarrow (\tilde{x}_i^* - \tilde{x}_{i+n}^*),\ 1 \leq i \leq N$}

3.2 Polytope Faces Pursuit with directional updates

As mentioned in [1, 76], the most computationally expensive operations of the PFP algorithm are the dictionary analysis computations $\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1}$ and $\tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}$ and the Moore-Penrose pseudo-inverse $(\tilde{\mathbf{A}}^k)^\dagger$ calculation required for the update of the solution vector $\tilde{\mathbf{x}}^k$ and its corresponding $\mathbf{c}^k$ at each iteration. For the pseudo-inverse estimation, greedy algorithms such as OMP usually adopt a Cholesky factorization method to update the matrix. According to this method, which has also been used in the original PFP implementation, if at step k of the algorithm $\tilde{\mathbf{A}}^k$ is a full column rank matrix then its pseudo-inverse is given by $(\tilde{\mathbf{A}}^k)^\dagger = [(\tilde{\mathbf{A}}^k)^T \tilde{\mathbf{A}}^k]^{-1} (\tilde{\mathbf{A}}^k)^T = [(\mathbf{R}^k)^T \mathbf{R}^k]^{-1} (\tilde{\mathbf{A}}^k)^T$, where $\mathbf{R}^k$ is an upper triangular matrix [77]. Although this method can be very efficient for small scale problems it requires storage of the upper triangular matrix $\mathbf{R}^k$, which grows in size by one column at each step corresponding to the selected atom of the dictionary. This storage requirement can be undesirable for large scale problems, as it increases memory requirements that usually affect the convergence speed of the algorithm. To overcome

these Cholesky factorization limitations iterative directional updates for the estimation of the solution vector $\tilde{\mathbf{x}}^k$ and the corresponding $\mathbf{c}^k$ should be used instead. These methods have first been addressed by Blumensath et al. in [78, 79, 80] for greedy algorithms. In particular, the work in [79] proposes a new set of greedy algorithms based on OMP, which make use of gradient methods to update the solution vector.

At the k-th iteration, the Polytope Faces Pursuit algorithm needs to update the solution vectors $\tilde{\mathbf{x}}^k$ and $\mathbf{c}^k$, respectively. Put another way, the algorithm needs to find updated solutions for the following inverse problems:

$$\min_{\tilde{\mathbf{x}}^k} \|\mathbf{y} - \tilde{\mathbf{A}}^k \tilde{\mathbf{x}}^k\|_2 \quad \text{and} \quad (\tilde{\mathbf{A}}^k)^T \mathbf{c}^k = \mathbf{1}. \tag{3.8}$$

One way to solve the above systems of linear equations is to use the very well known method of conjugate gradients [81, 82], which can free the algorithm from the burden of the storage of large matrices. However, this method requires that the given system of linear equations is positive-definite and square. In the examined problem, $\tilde{\mathbf{A}}^k$ is non-square. This can be handled by attempting to minimize the cost function:

$$J(\tilde{\mathbf{x}}) = \frac{1}{2} \tilde{\mathbf{x}}^T (\tilde{\mathbf{A}}^k)^T \tilde{\mathbf{A}}^k \tilde{\mathbf{x}}^k - (\tilde{\mathbf{A}}^k \tilde{\mathbf{x}}^k)^T \mathbf{y} \tag{3.9}$$

which is equivalent to solving the symmetric positive definite system [77]:

$$\min_{\tilde{\mathbf{x}}^k} \|(\tilde{\mathbf{A}}^k)^T \mathbf{y}^k - (\tilde{\mathbf{A}}^k)^T \tilde{\mathbf{A}}^k \tilde{\mathbf{x}}^k\|_2. \tag{3.10}$$

Therefore, an approach based on directional updates can be adopted by calculating the new directions and then updating both $\tilde{\mathbf{x}}^k$ and $\mathbf{c}^k$ at the k-th iteration of the algorithm.

3.2.1 Conjugate Gradient Polytope Faces Pursuit

The Conjugate Gradient Polytope Faces Pursuit (CG-PFP) algorithm can now be derived. At the k-th step of the CG-PFP, the algorithm should have selected k atoms from the given overcomplete dictionary $\tilde{\mathbf{A}}$ and removed $0 \leqslant l < k$ atoms, corresponding to the negative entries of the nonnegative vector $\tilde{\mathbf{x}}^k$. In order to estimate the coefficients vector the system $\tilde{\mathbf{x}}^k = \tilde{\mathbf{A}}^{k\dagger} \mathbf{y}$ needs to be solved, where $\tilde{\mathbf{A}}^k$ is an $M \times (k-l)$ matrix. However, by selecting a new atom, the dimensionality of the system increases as does the size of

the matrix defining the system. This increased dimensionality destroys the conjugacy of the previous directions and therefore at the k-th step of the algorithm a full conjugate gradient update is required.

The stopping criterion for the full conjugate gradient update is then the number of iterations, which is chosen to be as many as the size of the subset, namely the number of selected atoms. Alternatively, an error threshold on the objective function can be used as the stopping condition. The conjugate gradient update at the first iteration is:

$$\mathbf{d}_0^k = (\tilde{\mathbf{A}}^k)^T \mathbf{r}^{k-1} \tag{3.11}$$

$$\tilde{\mathbf{x}}_0^k = \tilde{\mathbf{x}}^{k-1} + \eta_0^k \mathbf{d}_0^k \tag{3.12}$$

$$\mathbf{r}_0^k = \mathbf{r}^{k-1} - \eta_0^k \tilde{\mathbf{A}}^k \mathbf{d}_0^k \tag{3.13}$$

where the step size η_0^k is given:

$$\eta_0^k = \frac{(\mathbf{r}^{k-1})^T \tilde{\mathbf{A}}^k \mathbf{d}_0^k}{(\tilde{\mathbf{A}}^k \mathbf{d}_0^k)^T \tilde{\mathbf{A}}^k \mathbf{d}_0^k}. \tag{3.14}$$

For the next iterations the update formulae becomes:

$$\mathbf{d}_i^k = (\tilde{\mathbf{A}}^k)^T \mathbf{r}_{i-1}^k + \beta_{i-1}^k \mathbf{d}_{i-1}^k \tag{3.15}$$

$$\tilde{\mathbf{x}}_i^k = \tilde{\mathbf{x}}_{i-1}^k + \eta_i^k \mathbf{d}_i^k \tag{3.16}$$

$$\mathbf{r}_i^k = \mathbf{r}_{i-1}^k - \eta_i^k \tilde{\mathbf{A}}^k \mathbf{d}_i^k \tag{3.17}$$

where the subscript i refers to the internal conjugate gradient iteration at the k-th step of the PFP algorithm. The step size η_i^k and the Gram-Schmidt constant β_i^k are given:

$$\eta_i^k = \frac{(\mathbf{r}_{i-1}^k)^T \tilde{\mathbf{A}}^k \mathbf{d}_i^k}{(\tilde{\mathbf{A}}^k \mathbf{d}_i^k)^T \tilde{\mathbf{A}}^k \mathbf{d}_i^k} \tag{3.18}$$

$$\beta_i^k = -\frac{(\tilde{\mathbf{A}}^k (\tilde{\mathbf{A}}^k)^T \mathbf{r}_i^k)^T \tilde{\mathbf{A}}^k \mathbf{d}_i^k}{(\tilde{\mathbf{A}}^k \mathbf{d}_i^k)^T \tilde{\mathbf{A}}^k \mathbf{d}_i^k}. \tag{3.19}$$

In order to estimate the corresponding $\mathbf{c}^k$ in (3.8) for the current solution vector $\tilde{\mathbf{x}}^k$ instead of another conjugate gradient update the algorithm follows a different strategy, which is summarized in the following theorem.

Theorem 5. *The basis vertex* $\mathbf{c}$ *at step* k *can be estimated using the iterative formula*

$$\mathbf{c}^k = \mathbf{c}^{k-1} + \lambda_k(\mathbf{r}^{k-1} - \mathbf{r}^k), \tag{3.20}$$

where λ_k *is the inverse atom selection criterion given by*

$$\lambda_k = \arg\min_{i \notin \mathcal{I}^{k-1}} \frac{1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}}{\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1}}. \tag{3.21}$$

Proof. Convenient update formulae for several of the quantities involved in the PFP algorithm can be calculated, including the quantities $\mathbf{B} \triangleq \tilde{\mathbf{A}}^\dagger = (\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^T$ and $P_{\tilde{\mathbf{A}}} \triangleq \tilde{\mathbf{A}}\tilde{\mathbf{A}}^\dagger = \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^T$. Using $\tilde{\mathbf{A}}^k = \left[\begin{array}{c|c} \tilde{\mathbf{A}}^{k-1} & \tilde{\mathbf{a}}_{i^k} \end{array}\right]$, it is straightforward to show that

$$\mathbf{B}^k = \left[\begin{array}{c} \mathbf{B}^{k-1}(\mathbf{I} - \tilde{\mathbf{a}}_{i^k}\mathbf{b}_{i^k}^T) \\ \hline \mathbf{b}_{i^k}^T \end{array}\right] \tag{3.22}$$

where $\mathbf{b}_{i^k} = (\mathbf{I} - P_{\tilde{\mathbf{A}}^{k-1}})\tilde{\mathbf{a}}_{i^k}/|(\mathbf{I} - P_{\tilde{\mathbf{A}}^{k-1}})\tilde{\mathbf{a}}_{i^k}|^2$, for any $\tilde{\mathbf{a}}_{i^k}$ such that $(\mathbf{I} - P_{\tilde{\mathbf{A}}^{k-1}})\tilde{\mathbf{a}}_{i^k} \neq \mathbf{0}$. Furthermore, under the same condition, it can also be shown that $\mathbf{r}^k = \mathbf{r}^{k-1} - \mathbf{b}_{i^k}\tilde{\mathbf{a}}_{i^k}^T\mathbf{r}^{k-1} = (\mathbf{I} - \mathbf{b}_{i^k}\tilde{\mathbf{a}}_{i^k}^T)\mathbf{r}^{k-1}$ and $\mathbf{c}^k = \mathbf{c}^{k-1} + \mathbf{b}_{i^k}(1 - \tilde{\mathbf{a}}_{i^k}^T\mathbf{c}^{k-1})$. Therefore for the switching in point for $\tilde{\mathbf{a}}_{i^k}$, the current point on the path through the polytope is given by $\mathbf{h} = \mathbf{c}^k + \lambda_k\mathbf{r}^k = \mathbf{c}^{k-1} + \lambda_k\mathbf{r}^{k-1}$ and hence $\mathbf{c}^k - \mathbf{c}^{k-1} = \lambda_k(\mathbf{r}^{k-1} - \mathbf{r}^k)$. □

Using the iterative formula of equation (3.20) for the estimation of the basis vertex $\mathbf{c}^k$ the algorithm can gain additional computational savings, as there is no longer the need to update the vector $\mathbf{c}$ by solving the computationally expensive linear system $(\tilde{\mathbf{A}}^k)^T\mathbf{c} = \mathbf{1}$ i.e. find $(\tilde{\mathbf{A}}^{k\dagger})^T$. The resulting algorithm of Conjugate Gradient Polytopes Faces Pursuit is summarized in Algorithm 3.

Comparing now CG-PFP with the existing stepwise implementation of PFP (Algorithm 2), one can notice that the release of the retarding constraints has been omitted in the proposed implementation. In other words, the algorithm at iteration k will not check and remove the appropriate columns of the subset $\tilde{\mathbf{A}}^k$, which correspond to coefficients of the solution vector $\tilde{\mathbf{x}}^k$ that violate the nonnegativity constraint. Due to the iterative nature of the conjugate gradient method switching out of these atoms will add computational cost to the algorithm. Every time a release occurs the algorithm will have to initialize both the solution vector $\tilde{\mathbf{x}}^k$ and the residual $\mathbf{r}^k$ and reestimate the current

Algorithm 3 Conjugate Gradient Polytope Faces Pursuit

Input: $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_i]$, $\mathbf{y}$ {If required, set $\tilde{\mathbf{A}} \leftarrow [\mathbf{A}, -\mathbf{A}]$}
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize: $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \tilde{\mathbf{A}}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \tilde{\mathbf{x}}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next face}
    $k \leftarrow k + 1$
    Find face: $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \{(\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1})/(1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}) \mid \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > 0\}$
    $\lambda_k \leftarrow (1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1})/(\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1})$
    Add constraints: $\tilde{\mathbf{A}}^k \leftarrow [\tilde{\mathbf{A}}^{k-1},\ \tilde{\mathbf{a}}_{i^k}],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
    $[\tilde{\mathbf{x}}^k, \mathbf{r}^k] \leftarrow$ update with directional updates (conjugate gradients)
    $\mathbf{c}^k \leftarrow \mathbf{c}^{k-1} + \lambda_k(\mathbf{r}^{k-1} - \mathbf{r}^k)$
end while
Output: $\mathbf{c}^* = \mathbf{c}^k$, $\tilde{\mathbf{x}}^* \leftarrow \mathbf{0}$ + corresponding entries from $\tilde{\mathbf{x}}^k$
{If required, get $x_i^* \leftarrow (\tilde{x}_i^* - \tilde{x}_{i+n}^*),\ 1 \le i \le N$}

solution vector for the whole subset apart from the columns that have been switched out. This process should repeat until all the entries of the solution vector are nonnegative. According to the standard implementation of PFP this step is necessary only to provide exact sparse solutions. If omitted the algorithm searches for approximate solutions that in practice will not deviate much from the optimal BP solution. In such a case, the path of the PFP algorithm as discussed in Section 3.3 will be identical to the path followed by the Least Angle Regression (LARS) algorithm [52], which is an interesting geometrical approach to solve (3.1), based on the concept of following the path which is equiangular among all current selected atoms. Nevertheless, as mentioned in [83] empirical findings show that sign constraint releases appear to be infrequent in practice.

Thus, CG-PFP stands as an alternative to the PFP algorithm when large scale sparse recovery problems encountered that can approximate the BP solution. In case that no negative entries are encountered as will be shown in Section 3.3, the convergence of CG-PFP and PFP will be identical and hence the algorithm will perform BP.

3.2.2 Approximations with directional updates

Although the Conjugate Gradient Polytope Faces Pursuit algorithm, as presented above, reduces the memory requirements and increases the convergence speed of the original PFP, it requires a full conjugate gradient solver each time a new face is encountered. Therefore, suboptimal directional updates can be used resulting in even faster but less accurate algorithms.

One suboptimal approach could be if the conjugate gradient solver is forced to iterate

only once at each step of the algorithm. The conjugate gradient solver at the first iteration will move towards the direction of the residual of the previous step. This is the same direction that the steepest descent method will follow at the first iteration. However, a more sophisticated approach is to consider approximate conjugate directions. In that case, the solution vector is updated by estimating the new direction that will be "conjugate" to the direction of the previous step of the algorithm. Therefore, the direction, the solution vector and the corresponding residual at the k-th step of the algorithm will be:

$$\mathbf{d}^k = (\tilde{\mathbf{A}}^k)^T \mathbf{r}^{k-1} + \beta^{k-1}\mathbf{d}^{k-1} \tag{3.23}$$

$$\tilde{\mathbf{x}}^k = \tilde{\mathbf{x}}^{k-1} + \eta^k \mathbf{d}^k \tag{3.24}$$

$$\mathbf{r}^k = \mathbf{r}^{k-1} - \eta^k \tilde{\mathbf{A}}^k \mathbf{d}^k. \tag{3.25}$$

where

$$\eta^k = \frac{(\mathbf{r}^{k-1})^T \tilde{\mathbf{A}}^k \mathbf{d}^k}{(\tilde{\mathbf{A}}^k \mathbf{d}^k)^T \tilde{\mathbf{A}}^k \mathbf{d}^k} \tag{3.26}$$

$$\beta^k = -\frac{(\tilde{\mathbf{A}}^k (\tilde{\mathbf{A}}^k)^T \mathbf{r}^k)^T \tilde{\mathbf{A}}^k \mathbf{d}^k}{(\tilde{\mathbf{A}}^k \mathbf{d}^k)^T \tilde{\mathbf{A}}^k \mathbf{d}^k}. \tag{3.27}$$

The resulting algorithm is called Approximate Conjugate Gradient Polytope Faces Pursuit (ACG-PFP).

3.3 Theoretical properties

This section provides the theoretical analysis of the PFP algorithm. After the presentation of the necessary and sufficient recovery conditions that guarantee uniqueness of the optimum solution pair ($\tilde{\mathbf{x}}^*$, $\mathbf{c}^*$), the relation between the Fuchs Condition [59] and the Exact Recovery Condition (ERC) [35] is investigated and it is shown that when ERC is satisfied the PFP and CG-PFP algorithms will be equivalent in the sense that they will both yield the optimal sparse solution following the exact same path. Empirical findings verify the theoretical results.

3.3.1 Fuchs Condition & ERC

Suppose that condition (3.5a) holds for some $\tilde{\mathbf{x}}_0$ with $\|\tilde{\mathbf{x}}_0\|_0 \leq K$ nonzero elements and define $\tilde{S} = \text{supp}(\tilde{\mathbf{x}}_0)$ to be the set of indices corresponding to the nonzero entries of the nonnegative vector $\tilde{\mathbf{x}}_0$. Then the following result gives the necessary and sufficient conditions for $\tilde{\mathbf{x}}_0$ to be the unique optimum of (3.2).

Theorem 6. (Plumbley [73]) *$\tilde{\mathbf{x}}_0$ is the unique optimum point of the primal linear program if and only if a) $\tilde{\mathbf{A}}_{\tilde{S}}$ has full rank and b) there exists some $\mathbf{c} \in \mathbb{R}^M$ such that*

$$\tilde{\mathbf{a}}_j^T \mathbf{c} = 1, \quad \forall j \in \tilde{S} \tag{3.28a}$$

$$\tilde{\mathbf{a}}_j^T \mathbf{c} < 1, \quad \forall j \notin \tilde{S}. \tag{3.28b}$$

Theorem 6 is the equivalent to the Fuchs Condition [59] for the problem (3.1) in its standard form. Therefore, Theorem 6 guarantees ℓ_1-unique-optimality and in addition it provides the weakest possible conditions for ℓ_1-unique-optimality. However, it is rather hard to test given an underdetermined system of linear equations. As showed in [73], a stronger condition than necessary for ℓ_1-uniqueness holds, which is an extension of the Fuchs Corollary [84] for the original form of problem (3.1) to its standard form.

Corollary 1. (Plumbley [73]) *Let $\tilde{\mathbf{x}}_0$ be a solution to the primal linear program of equation* (3.2)*. If $\tilde{\mathbf{A}}_{\tilde{S}}$ has full rank and*

$$\tilde{\mathbf{a}}_j^T \mathbf{c} < 1, \quad \forall j \notin \tilde{S} \tag{3.29}$$

is satisfied with the dual vector $\mathbf{c} = (\tilde{\mathbf{A}}_S^{\dagger})^T \mathbf{1}$, then $\tilde{\mathbf{x}}_0$ is the unique optimum.

Nevertheless, a more well known recovery condition than the Fuchs Condition is the Exact Recovery Condition (2.44) of Theorem 3, presented in Section 2.2.4. Interestingly, as pointed out by Gribonval and Nielsen [58], Fuchs Condition and ERC are closely related to each other. More specifically, by rewriting condition (3.29) of Corollary 1 to its original form, the inequality becomes $|\mathbf{a}_j^T (\mathbf{A}_S^{\dagger})^T \text{sgn}(\mathbf{x}_S)| < 1, \ \forall j \notin S$, where $S = \text{supp}(x)$

and sgn(·) denotes the signum function defined as:

$$\text{sgn}(x) = \begin{cases} -1 & \text{if} \quad x < 0 \\ 0 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x < 0 \end{cases} . \tag{3.30}$$

Therefore, it holds:

$$\begin{aligned} &\max_{\mathbf{x}_S} \max_{j \notin S} |\mathbf{a}_j^T (\mathbf{A}_S^\dagger)^T \text{sgn}(\mathbf{x}_S)| \\ &= \max_{\mathbf{x}_S} \max_{j \notin S} |\text{sgn}(\mathbf{x}_S^T) \mathbf{A}_S^\dagger \mathbf{a}_j| \\ &= \max_{\mathbf{x}_S} \max_{j \notin S} |\langle \text{sgn}(\mathbf{x}_S), \mathbf{A}_S^\dagger \mathbf{a}_j \rangle| \\ &= \max_{j \notin S} |\langle \mathbf{1}, \mathbf{A}_S^\dagger \mathbf{a}_j \rangle| \\ &= \max_{j \notin S} \|\mathbf{A}_S^\dagger \mathbf{a}_j\|_1 < 1. \end{aligned} \tag{3.31}$$

Consequently, the ERC of Theorem 3 can be viewed as a corollary to Corollary 1, which therefore makes it a stronger condition than both Fuchs Condition and Fuchs Corollary. In other words, if ERC is met then the conditions in Theorem 6 and Corollary 1 will both be satisfied. It is clear though that the opposite argument does not always hold.

For the original version of PFP which considers removal of atoms, Corollary 1 is sufficient for ℓ_1 uniqueness [73]. The main result presented here shows that when ERC is also satisfied then PFP will converge to the optimum solution $\tilde{\mathbf{x}}^*$ in only K steps. This is stated more formally in the following theorem.

Theorem 7. *Let* $\tilde{\mathbf{x}}^*$ *and* $\mathbf{c}^*$ *be the optimum solution pair to the primal and dual linear programs of* (3.2) *and* (3.4) *respectively and define* $\mathbf{A}_S$ *as in theorem* (3)*. If*

$$\max_{j \notin S} \|\mathbf{A}_S^\dagger \mathbf{a}_j\|_1 < 1 \tag{3.32}$$

then PFP will find the optimum solution pair $\tilde{\mathbf{x}}^*$ *and* $\mathbf{c}^*$ *in* K *steps.*

Proof. In order to prove this, it is required to show that the algorithm at the k-th iteration will select a dictionary atom $\tilde{\mathbf{a}}_j$ so that $j \in \tilde{S}$ and that no atom removals will occur in

any of the first K iterations. For the first part, note that at the k-th iteration and for $j \in \tilde{S}$, it holds:

$$\begin{aligned}
\arg\max_{j \notin \mathcal{I}^{k-1}} \frac{\tilde{\mathbf{a}}_j^T \mathbf{r}^{k-1}}{1 - \tilde{\mathbf{a}}_j^T \mathbf{c}^{k-1}} &\geq \frac{\sum_{i=k}^{K} x_k(i) \tilde{\mathbf{a}}_j^T \tilde{\mathbf{a}}_i}{1 - \tilde{\mathbf{a}}_j^T [(\tilde{\mathbf{A}}^{k-1})^\dagger]^T \mathbf{1}} \\
&\geq \frac{x_k(j) - \sum_{i \neq j} x_k(i) \tilde{\mathbf{a}}_j^T \tilde{\mathbf{a}}_i}{1 + \| \tilde{\mathbf{a}}_j^T [(\tilde{\mathbf{A}}^{k-1})^\dagger]^T \|_1} \\
&\geq \frac{x_k(j) - \mu \sum_{i \neq j} x_k(i)}{1 + \| [(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{A}}^{k-1}]^{-1} \|_1 \| (\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{a}}_j \|_1} \\
&\geq \frac{x_k(j) - \mu(K-k) x_k(j)}{1 + \mu(k-1)/(1 - \mu(k-2))}
\end{aligned} \tag{3.33}$$

where $\mathcal{I}^{k-1}$ is the set of indices corresponding to the already selected atoms and μ is the dictionary mutual coherence. In the above inequalities, it has been used the fact that the residual at the k-th iteration of PFP can be expressed as a linear combination of $K-k$ atoms that have not been yet selected and belong to the true support set. Hence, it is assumed that at the previous $k-1$ iterations the algorithm has selected atoms from the support set i.e. $\mathcal{I}^{k-1} \subset \tilde{S}$. In that case, the residual will be orthogonal to the $k-1$ selected atoms, as the algorithm by definition enforces this orthogonality. For the denominator, in the last inequality of (3.33), it has been used the fact that according to the Neumann series for any matrix $\mathbf{D}$ such that $\|\mathbf{D}\|_1 < 1$, it holds:

$$\sum_{l=0}^{\infty} \mathbf{D}^l = (\mathbf{I} - \mathbf{D})^{-1}. \tag{3.34}$$

Notice that the Gramian $(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{A}}^{k-1}$ has ones in its diagonal. It can therefore be expressed as $(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{A}}^{k-1} = \mathbf{I} + \mathbf{D}$. Using now (3.34), it follows:

$$\begin{aligned}
\| [(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{A}}^{k-1}]^{-1} \|_1 &= \| (\mathbf{I} + \mathbf{D})^{-1} \|_1 \\
&= \left\| \sum_{l=0}^{\infty} -\mathbf{D}^l \right\|_1 \leq \sum_{l=0}^{\infty} \|\mathbf{D}\|_1^l \\
&= \frac{1}{1 - \|\mathbf{D}\|_1} \leq \frac{1}{1 - \mu(k-2)}.
\end{aligned} \tag{3.35}$$

On the other hand for dictionary atoms $\tilde{\mathbf{a}}_j$ with $j \notin \tilde{S}$, the algorithm requires:

$$\begin{aligned}
\arg\max_{j \notin \mathcal{I}^{k-1}} \frac{\tilde{\mathbf{a}}_j^T \mathbf{r}^{k-1}}{1 - \tilde{\mathbf{a}}_j^T \mathbf{c}^{k-1}} &\leq \frac{\sum_{i=k}^{K} x_k(i) \tilde{\mathbf{a}}_j^T \tilde{\mathbf{a}}_i}{1 - \tilde{\mathbf{a}}_j^T [(\tilde{\mathbf{A}}^{k-1})^\dagger]^T \mathbf{1}} \\
&\leq \frac{\sum_{i=k}^{K} x_k(i) \tilde{\mathbf{a}}_j^T \tilde{\mathbf{a}}_i}{1 - \|\tilde{\mathbf{a}}_j^T [(\tilde{\mathbf{A}}^{k-1})^\dagger]^T\|_1} \\
&\leq \frac{\mu \sum_{i=k}^{K} x_k(i)}{1 - \|[(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{A}}^{k-1}]^{-1}\|_1 \|(\tilde{\mathbf{A}}^{k-1})^T \tilde{\mathbf{a}}_j\|_1} \\
&\leq \frac{\mu(K-k+1) x_k(j)}{1 - \mu(k-1)/(1-\mu(k-2))}.
\end{aligned} \tag{3.36}$$

The combination of equations (3.33) and (3.36) yields:

$$2\mu K < 1 - \mu + 2\mu k - \frac{\mu(k-1)}{1-\mu(k-2)}(1+\mu). \tag{3.37}$$

Minimizing the quantity $q(k) = 2\mu k - \frac{\mu(k-1)}{1-\mu(k-2)}(1+\mu)$ with respect to k can further bound the sparsity level K. Considering that PFP requires $|\tilde{\mathbf{a}}_j^T \mathbf{c}^{k-1}| < 1$ and therefore $\frac{\mu(k-1)}{1-\mu(k-2)} < 1$, it is straightforward to show that $q(k)$ will take its minimum value at $k = 1$. Consequently, substituting $k = 1$, equation (3.37) yields:

$$K < \frac{(1+\mu)}{2\mu} \Rightarrow K < \frac{\mu^{-1}+1}{2}. \tag{3.38}$$

The inequality (3.38) implies the ERC. According to Theorem 4 of Section 2.2.4, if ERC is satisfied then the threshold on the sparsity level K is given by the above inequality.

Therefore, if (3.38) holds then PFP at the k-th iteration will select an atom from the true support set $\tilde{S}$. It remains to be seen that this condition can guarantee that no atom removals will occur during the first K iterations. To do this, once again it is assumed that at the k-th iteration PFP has selected $k-1$ atoms from $\tilde{S}$ without any removals. Note that at any iteration it holds:

$$\hat{\mathbf{y}}_k^T \mathbf{c}^* = \hat{\mathbf{x}}_k^T (\tilde{\mathbf{A}}^k)^T (\tilde{\mathbf{A}}_{\tilde{S}}^\dagger)^T \mathbf{1} \leq \hat{\mathbf{x}}_k^T \mathbf{1} \leq \|\hat{\mathbf{x}}_k\|_1 \tag{3.39}$$

where $\hat{\mathbf{y}}_k$ denotes the estimate of $\mathbf{y}$ at the end of the k-th iteration. It is clear that if (3.32) holds for the given system, then the inequality (3.39) will hold only if $\tilde{\mathbf{A}}^k$ contains dictionary atoms $\tilde{\mathbf{a}}_j$ with $j \notin S$. However, this is contradictory according to (3.38) and

thus if (3.32) is satisfied the equality in (3.39) will hold. Subsequently, when equation (3.32) holds, PFP at the k-th iteration will select an atom from $\tilde{S}$ and also the nonnegative constraint on $\tilde{\mathbf{x}}$ cannot be violated. It follows that the algorithm will find the optimum solution pair in $\tilde{\mathbf{x}}^*$ and $\mathbf{c}^*$ in K iterations. □

The result of Theorem 7 has also been observed in [33] for the Homotopy method, which differs from LARS algorithm by allowing atom removals at each stage. The authors eventually show that Homotopy and LARS will follow the same path when ERC is met and recover the sparse solution in K steps. From that point of view, if the atom removal step is omitted from PFP then its path will be identical to the LARS path. The same holds for CG-PFP algorithm which by definition does not consider switching out of bases. In the case now that ERC holds for the given overcomplete dictionary $\mathbf{A}$, then an outcome of Theorem 7 is the following corollary.

Corollary 2. *Consider $\mathbf{A}_S$ as in Theorem 3. Then if*

$$\arg\max_{j\notin S} \|\mathbf{A}_S^{\dagger}\mathbf{a}_j\|_1 < 1 \tag{3.40}$$

PFP and CG-PFP algorithms will be equivalent. Both algorithms will find the optimum solution pair $\tilde{\mathbf{x}}^*$ and $\mathbf{c}^*$ in K steps following the exact same path.

Proof. It follows naturally from Theorem 7. □

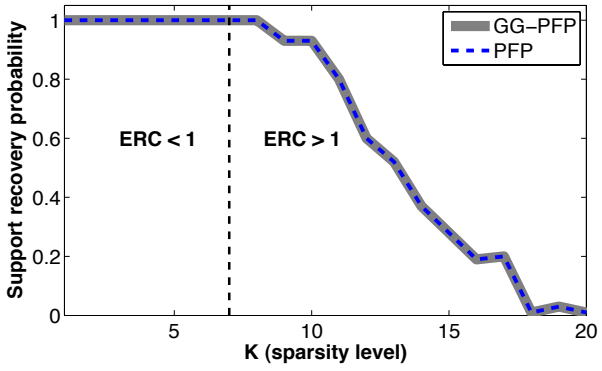


Figure 3.3: Support recovery rates (over 100 trials) of the K-term approximation of PFP and CG-PFP vs the sparsity level K for a dictionary $\mathbf{A} \in \mathbb{R}^{M\times N}$ with $M = 128$ and $N = 256$. The dotted line separates the region of sparsity levels K for which ERC is satisfied from the region where ERC is not satisfied.

The result of Corollary 2 was also verified empirically. This was done by comparing

the K-term approximation of PFP and CG-PFP algorithms, where both algorithms were forced to terminate after performing K iterations. The generated dictionary was chosen to be twice overcomplete $\mathbf{A} \in \mathbb{R}^{128\times256}$ and its columns were drawn from an i.i.d. zero-mean Gaussian distribution. The sparse vectors $\mathbf{x}$ were generated by drawing K elements from an i.i.d. process and placing them at K random entries of $\mathbf{x}$. The average full support recovery probabilities over 100 iterations are plotted in Fig. 3.3. As can be seen, as long as the ERC is satisfied, the rate of full support recovery is 1 for both algorithms, whereas for larger sparsity levels such that $K \geq 8$, ERC is not satisfied and as expected both algorithms do not converge to the optimum solution in K steps. Hence, these empirical findings demonstrate the theoretical results of Theorem 7 and Corollary 2.

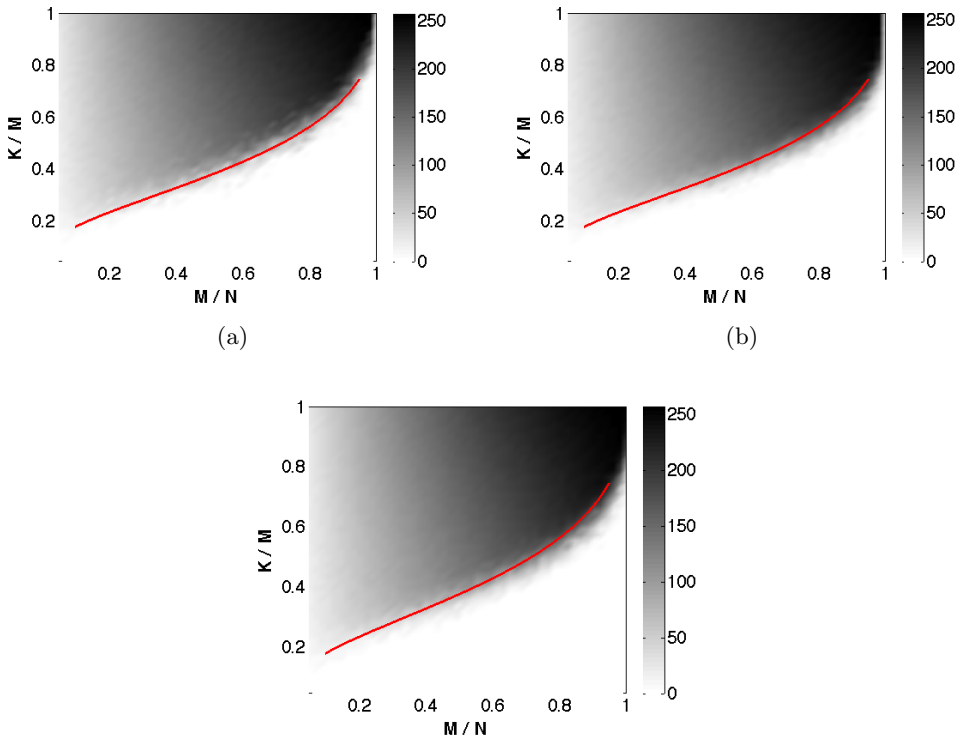
However, the above performance cannot be guaranteed for both algorithms when only Fuchs Corollary holds and ERC is not satisfied. This is demonstrated in the following section using an average case analysis for both algorithms.

3.3.2 Phase transitions of PFP algorithms.

Phase diagrams have been used extensively in order to quantify the performance of compressed sensing and sparse recovery algorithms. Initially introduced by Donoho and Tanner [85], these phase diagrams depict the performance of an algorithm as a function of the over-sampling ratio M/N and the under-sampling ratio K/M. Both of these quantities are chosen to vary within the range [0, 1]. According to the selected performance measure (e.g. Euclidean distance between the recovered and the true sparse vector $\mathbf{x}$) phase diagrams display a phase transition [33]. This phase transition phenomenon can give more insight on the behaviour of an algorithm between success and failure in recovering $\mathbf{x}$.

Fig. 3.4 shows the phase diagrams of the original PFP algorithm with and without atom removals and the proposed CG-PFP. The red curve corresponds to the theoretical curve for ℓ_1 minimization. The shaded area is the number of coefficients that differ more than 10^{-4} from the optimal sparse solution. Therefore, the darker the region the lower the recovery capabilities of the algorithm. The results shown have been obtained using dictionaries drawn from the random uniform ensemble and the number of dictionary atoms is $N = 256$.

The phase transitions of the three diagrams, obtained by these average case results,



(a) (b)

(c)

Figure 3.4: Phase diagrams of (a) orginal PFP algorithm (b) original PFP algorithm without atom removals and (c) CG-PFP algorithm.

verify the theoretical results presented in the Section 3.3.1. As expected, for small values of K/M and for the whole range of the ratio M/N the recovery performance is identical for all algorithms. However, for larger values of K/M, where ERC is less likely to be satisfied the original version exhibits slightly better performance.

3.4 Stagewise Polytope Faces Pursuit

Stagewise greedy algorithms such as Stagewise Orthogonal Matching Pursuit (StOMP) [86] and Stagewise Weak Orthogonal Matching Pursuit (SWOMP) [80] attempt to add several dictionary atoms instead of one atom at a time. As a result, these algorithms provide approximate solutions but they achieve faster convergence. In a similar manner to the aforementioned stagewise versions of OMP, in [76] the authors proposed the Stagewise Polytope Faces Pursuit (StPFP) algorithm, which allows for a variable number of selected basis vectors at each stage. More specifically, once the adjusted correlations have been calculated, the algorithm at the k-th iteration selects $q^k \geq 1$ atoms according to a basis selection criterion.

The basis selection criterion could be either a fixed selection criterion or the weak selection criterion, which is a modification of the weak-MP [28] selection criterion. The fixed selection criterion is straightforward; the algorithm at stage k will select the $q^k = q$ atoms with the largest scaled correlation. A more sophisticated approach is to use the weak selection criterion, similar to the work of Davies and Blumensath for the SWOMP algorithm [80]. According to this approach, the algorithm at the k-th stage first estimates the scaled correlation:

$$\theta(\tilde{\mathbf{a}}_i^k) = \frac{\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1}}{1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}} \tag{3.41}$$

for all atoms that have not yet been selected by the algorithm. The weak selection criterion is then formulated as:

$$|\theta(\tilde{\mathbf{a}}_i^k)| \geq \beta \max_i |\theta(\tilde{\mathbf{a}}_i^k)| \tag{3.42}$$

where $0 < \beta \leq 1$ is a threshold control parameter. Therefore, the weak selection criterion applies a threshold relative to the maximum of the scaled correlations.

After the atoms have been selected, the algorithm updates the solution vectors and the residual and iterates until the stopping conditions are met, similar to Algorithm 2.

3.4.1 Stagewise Conjugate Gradient Polytope Faces Pursuit

This section presents the stagewise version of the algorithm using directional updates. Once q^k new atoms have been identified by the fixed or weak selection criterion, the full conjugate gradient method is used to update the solution vector $\tilde{\mathbf{x}}^k$, in a similar way to that discussed in Section 3.2. The main algorithmic difference between the stepwise and stagewise versions appears when it comes to the estimation of the corresponding $\mathbf{c}^k$ to the current solution vector. This is because the iterative update formula of equation (3.20) cannot be used in that case, due to the fact that the update for $\mathbf{B}^k$ of equation (3.22) will not hold when adding $q^k > 1$ dictionary atoms. Subsequently, a second directional update strategy for the estimation of $\mathbf{c}^k$ is considered.

At the k-th stage, the algorithm needs to solve the system for $\mathbf{c}^k$:

$$(\tilde{\mathbf{A}}^k)^T \mathbf{c}^k = \mathbf{1}^k \tag{3.43}$$

Algorithm 4 Stagewise Polytope Faces Pursuit with directional updates

Input: $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_i]$, $\mathbf{y}$ {If required, set $\tilde{\mathbf{A}} \leftarrow [\mathbf{A}, -\mathbf{A}]$}
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize: $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \tilde{\mathbf{A}}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \tilde{\mathbf{x}}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next face}
    $k \leftarrow k + 1$
    $q^k \leftarrow$ select a suitable number of vectors to be added
    Find q^k faces:
    $\mathcal{J}^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}}^{q^k} \{(\tilde{\mathbf{a}}_i^T \mathbf{r}^{k-1})/(1 - \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1}) \mid \tilde{\mathbf{a}}_i^T \mathbf{c}^{k-1} > 0\}$
    Add constraints: $\tilde{\mathbf{A}}^k \leftarrow [\tilde{\mathbf{A}}^{k-1},\ \{\tilde{\mathbf{a}}_{i^k} \mid i \in \mathcal{J}^k\}],\ \mathcal{I}^k \leftarrow \mathcal{J}^{k-1} \cup \mathcal{J}^k$
    $[\tilde{\mathbf{x}}^k, \mathbf{c}^k, \mathbf{r}^k] \leftarrow$ update with directional updates
end while
Output: $\mathbf{c}^* = \mathbf{c}^k$, $\tilde{\mathbf{x}}^* \leftarrow \mathbf{0} +$ corresponding entries from $\tilde{\mathbf{x}}^k$
{If required, get $x_i^* \leftarrow (\tilde{x}_i^* - \tilde{x}_{i+n}^*),\ 1 \leq i \leq N$}

where the dimensionality of vector $\mathbf{1}^k$ increases at each stage by as much as the total number of the newly selected atoms. Vector $\mathbf{c}^k$ is estimated using the steepest descent method with a number of iterations equal to the size of the current selected support set, as in practice empirical results suggest that the accuracy in the estimation does not have a large effect in the atom selection criterion. Put another way, an approximate estimate of $\mathbf{c}^k$ is often enough for the algorithm to select the correct atoms at each stage. The Stagewise Conjugate Gradient Polytopes Faces Pursuit (CG-StPFP) is illustrated in Algorithm 4.

3.5 Experimental results

For the evaluation of the performance of the proposed methods of stepwise and stagewise PFPs, the algorithms were compared against state-of-the-art algorithms, such as Basis Pursuit and Gradient Projections for Sparse Reconstruction (GPSR) algorithm. First of all, all algorithms were tested on synthetic data and then on problems generated using the benchmarking MATLAB toolbox SMALLBOX [87, 88], including a compressed sensing and an audio source separation problem.

3.5.1 Experimental evaluation on synthetic data with fixed problem size

For the first set of experiments, a sparse representation problem was generated using synthetic data. More specifically, the dictionary $\mathbf{A} \in \mathbb{R}^{128 \times 256}$ was a two times redundant matrix with normalized columns, generated by drawing its columns from an i.i.d. normal

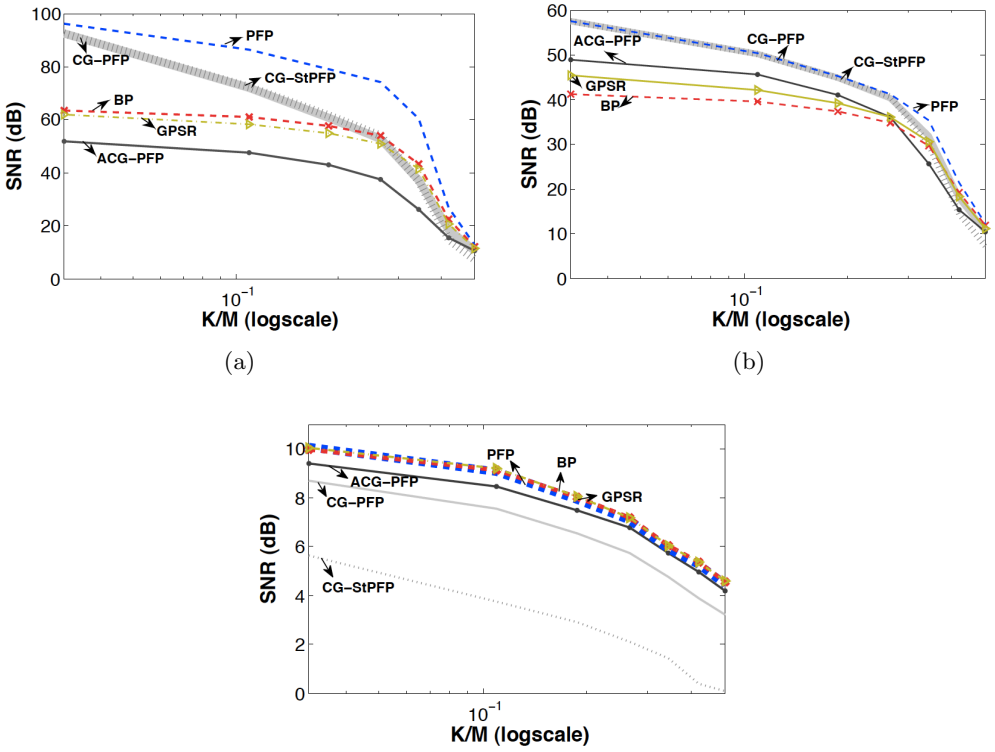


Figure 3.5: Average SNR values of the estimated sparse coefficient vectors for variable sparsity levels for the PFP, BP, GPSR, CG-PFP, ACG-PFP and CG-StPFP algorithms. The observations have been corrupted by additive Gaussian noise resulting in SNR of (a) 100 dB, (b) 60 dB and (c) 30 dB.

distribution. The entries of the sparse vector $\mathbf{x} \in \mathbb{R}^{256}$ were also drawn from an i.i.d. normal distribution and the sparsity level K varied from 4 to 64 with a step of size 10. Next, the observations were generated as $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$, where $\mathbf{n} \sim \mathcal{N}(0, \sigma^2)$ is the i.i.d Gaussian additive noise with variable variance σ^2 resulting in different SNR values (100 dB, 60 dB and 30 dB).

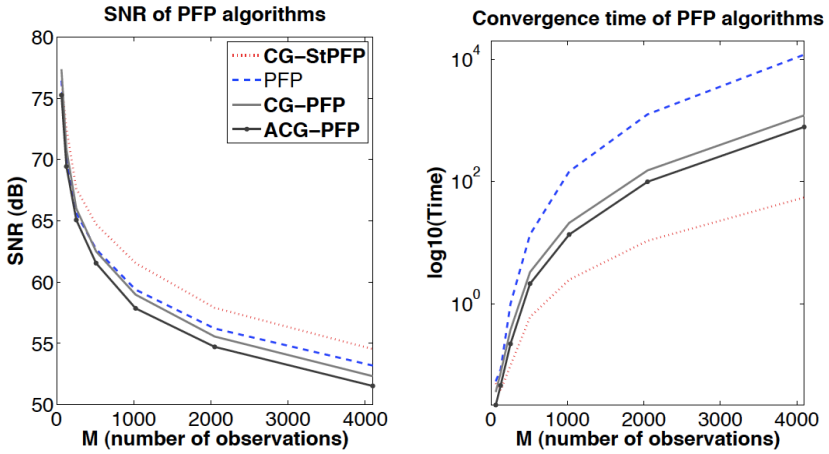
The proposed stepwise and stagewise implementations of PFP based on directional updates were compared against the SparseLab implementation of Basis Pursuit (SolveBP) [89, 90] and the GPSR algorithm [53] taken from the Gpsr toolbox (GPSR_Basic) [91]. Both toolboxes are included in the SmallBox framework. Parameters, such as number of iterations or stopping error criterion were left at their default values. After the estimation of the recovered signal for each algorithm, the SNR values were computed and the results were plotted against the sparsity level (ratio between nonzero entries K and the dimension of the observations M). All results have been averaged over 1000 iterations.

Fig. 3.5(a)-(c) illustrate the results of the stepwise CG-PFP and ACG-PFP algorithms

(PFP based on the full conjugate gradient directional update and on the approximate conjugate gradient update respectively), the stepwise PFP (PFP based on Cholesky factorization), the stagewise CG-StPFP algorithm under the weak selection criterion, BP and GPSR at the three different noise levels respectively. The parameter β was set at 0.8 for the weak selection criterion. As can be seen in Fig. 3.5(a), for the sparsity level $K = 14$ and therefore $K/M \approx 0.11$, PFP achieves the best performance resulting in SNR = 86.5 dB followed by CG-PFP and CG-StPFP, which both achieve an SNR of approximately 72 dB. All three algorithms outperform BP and GPSR, for which the obtained SNR is 61 dB and 58 dB, respectively. However, the approximate PFP algorithm exhibits the worst performance with SNR at 47 dB. In the case when the additive noise in the measurements results in SNR of 60 dB (Fig. 3.5(b)), it can be observed that for the same ratio of K/M all PFP algorithms perform better than BP and GPSR. In more detail, the SNR values for PFP, CG-PFP and CG-StPFP is approximately 50 dB and 46 dB for the ACG-PFP algorithm. On the other hand, the performance falls at 39 dB and 42 dB for BP and GPSR respectively. Finally, at the same sparsity level $K = 14$ Fig. 3.5(c) (SNR = 30 dB) shows that PFP, BP and GPSR perform close to 9 dB, while CG-PFP and ACG-PFP achieve SNR values of 8.5 dB and 7.5 dB, respectively. In that case, the performance of CG-StPFP degrades dramatically, as it drops at only 4 dB. In general, it could be said that the Cholesky factorization PFP method shows a slightly better overall performance than the conjugate gradient based one. However, the CG-PFP algorithm performs quite close to PFP in all three cases. It is worth noting that although ACG-PFP displays the worst performance in the first case when the noise level is at 100 dB (Fig. 3.5(a)), its performance does not degrade as much as CG-PFP when the noise level increases. This might be due to the fact that in the noisy case less iterations of the conjugate gradient step or suboptimal directional updates behave better. On the other hand, CG-StPFP performs close to CG-PFP in the first two experiments (Fig. 3.5(a)-(b)), but when the SNR is at 30 dB the algorithm performs poor. This suggests that when the noise level is higher, choosing more atoms at a time does not necessarily lead to good approximations.

3.5.2 Experimental evaluation on synthetic data with variable problem size

In the second experiment, the PFP algorithms' performance for a variable problem size was compared. To do this, the number of observations M was increased from 64 to 4,096 while the number of variables was increased proportionally as $N = 2M$. The sparsity level K was kept constant at $K/M = 0.125$. The columns of the overcomplete dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ were drawn from an i.i.d. normal distribution. The sparse vector $\mathbf{x}$ was also generated by drawing K elements from an i.i.d. distribution and placing them at K random entries of $\mathbf{x}$. The observations vector was generated as $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ is the i.i.d. Gaussian additive noise. The SNR was kept constant at 90 dB for this experiment.



(a) (b)

Figure 3.6: (a) Average SNR values of the estimated sparse coefficient vectors and (b) average elapsed times for the CG-PFP, PFP, ACG-PFP and CG-StPFP (with the weak selection criterion) against the problem size.

Fig. 3.6(a) shows the SNR performance of the three stepwise PFP methods (original PFP, CG-PFP and ACG-PFP) and the stagewise based on the weak selection criterion (CG-StPFP) and Fig. 3.6(b) the elapsed times, respectively. Regarding the signal recovery, the stagewise algorithm achieves the best performance for this experiment, especially as the problem size increases. It therefore suggests that for large scale problems stagewise approaches can also improve the recovery. The stepwise algorithms perform closely and the one using the approximate directional updates shows the worst performance in terms of SNR but still close to the CG-PFP and PFP functions. Moreover, Fig. 3.6(b) clearly shows the advantage of using directional updates. Apart from the fact that storage re-

quirements are decreased significantly, the gain in speed becomes larger as the problem size increases. When the directional updates and the stagewise approach are combined the algorithm results in the fastest convergence. For example, when the number of measurements is $M = 2048$, it can be noted that CG-StPFP requires only 11 sec to yield a sparse approximation. This makes the algorithm approximately 9 times faster than AC-PFP, which converges after 99 sec and 14 times faster than CG-PFP, which converges after 152 sec. PFP is much slower in this case with the average time to convergence being approximately 1256 sec, which is 8 times slower than the average time of CG-PFP.

3.5.3 Evaluation on SPARCO problems

SPARCO [92, 93] is a MATLAB framework containing a suite of different signal processing problems, which makes it suitable for benchmarking sparse reconstruction algorithms. SPARCO is included in the SMALLBOX toolkit. A compressed sensing, an image deblurring and an audio source separation problem has been chosen to evaluate the proposed algorithms.

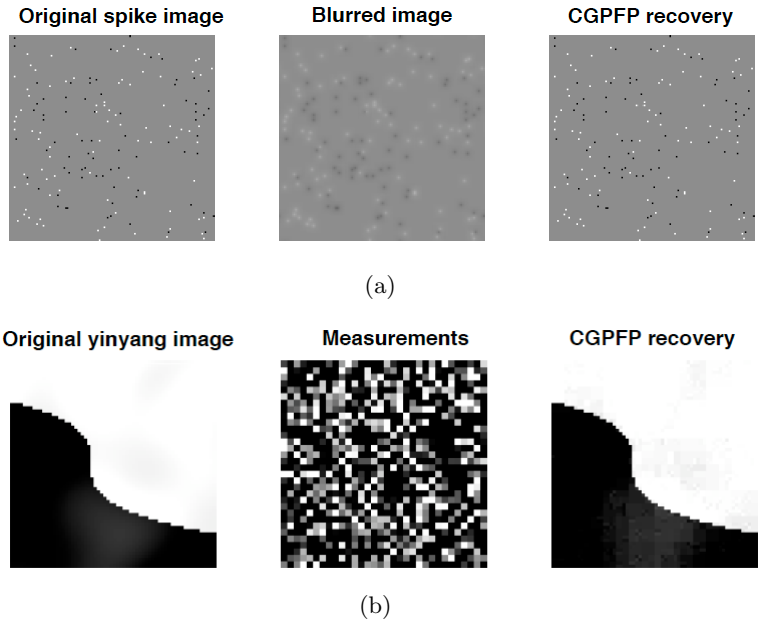


Figure 3.7: Original signal, observations and signal recovery by CG-PFP algorithm for (a) image deblurring SPARCO problem 702 and (b) 2D compressed sensing SPARCO problem 603.

SPARCO problem 702 is an image deblurring problem, consisting of a binary spike image. The signal is blurred by convolution with an 8×8 blurring mask and normally distributed noise with standard deviation 0.01 is added to the blurred signal. The prob-

Table 3.1: Signal-to-Nosie-Ratio and Elapsed times (problems 702 and 603)

	problem 702		problem 603	
	SNR (dB)	Time (sec)	SNR (dB)	Time (sec)
BP	17.11	227.24	30.16	287.7
GPSR	**19.19**	34.25	28.71	22.18
PFP	**19.11**	13.41	**30.81**	43.86
StPFP	**19.52**	8.08	**30.88**	18.77
CG-PFP	**19.11**	12.18	**30.96**	32.05
ACG-PFP	17.67	**5.1**	**30.9**	**15.88**
CG-StPFP	**19.24**	**5.7**	**30.83**	**7.54**

lem size is $M = 16384$, $N = 16384$ and the sparsity level is $K = 164$. Fig. 3.7(a) shows the original binary image, the blurred observations and the CG-PFP reconstruction. The stepwise versions of the Polytope Faces Pursuit algorithm (PFP, CG-PFP, ACG-PFP) and the stagewise algorithms (StPFP, CG-StPFP) under the fixed selection criterion were tested against the BP solver and the GPSR algorithm. For all experiments the number of selected atoms per iteration has been set to 10 for the fixed selection criterion and the parameter β at 0.8 for the weak selection criterion of the stagewise sparse recovery algorithms. Table 3.1 summarises the results in terms of SNR performance and convergence time for the tested algorithms.

As can be seen, most of the algorithms SNR performance is close to 19 dB. However, it is clear that the Polytope Faces Pursuit algorithms based on directional updates converge much faster, with the stagewise version of CG-PFP achieving the best performance considering both SNR and elapsed time. The SNR recovery of the approximate stepwise PFP algorithm deteriorates, but it converges much faster outperforming all other algorithms in terms of convergence speed.

Problem 603 is a 2D compressed sensing problem of the yinyang image shown in Fig. 3.7(b). The original signal size is 64×64 and it is measured using a $M = 1024 \times N = 4096$ binary ensemble with unit norm columns. The additive noise in the measurements is normally distributed with standard deviation 0.01. The original image is not exact sparse but compressible in the wavelet domain and a square transform matrix containing Daubechies atoms is used for the decomposition. Most of the compared algorithms perform similarly regarding the signal recovery. However, their convergence time varies significantly and once more, Table 3.1 encapsulates the gain of using directional updates, especially when combined with the stagewise approach.

Finally, the performance of the proposed framework was evaluated on an audio source

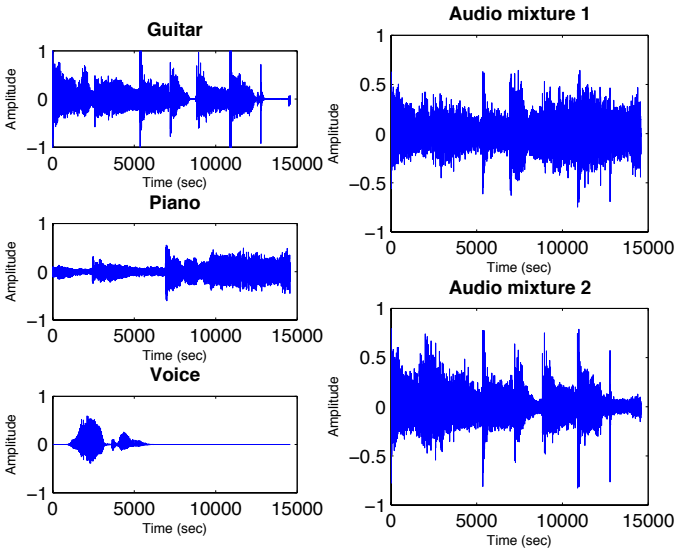


Figure 3.8: Source separation SPARCO problem 402, consisting of three audio sources: guitar, piano and voice (left) and two mixtures (right).

separation problem. SPARCO problem 402 is an underdetermined source separation problem consisting of three audio sources (guitar, piano and voice) and 2 mixtures (Fig. 3.8). The mixing is instantaneous and a windowed DCT transform is used to provide the sparse representations of the audio signals. All sources have been initially downsampled to 8 kHz.

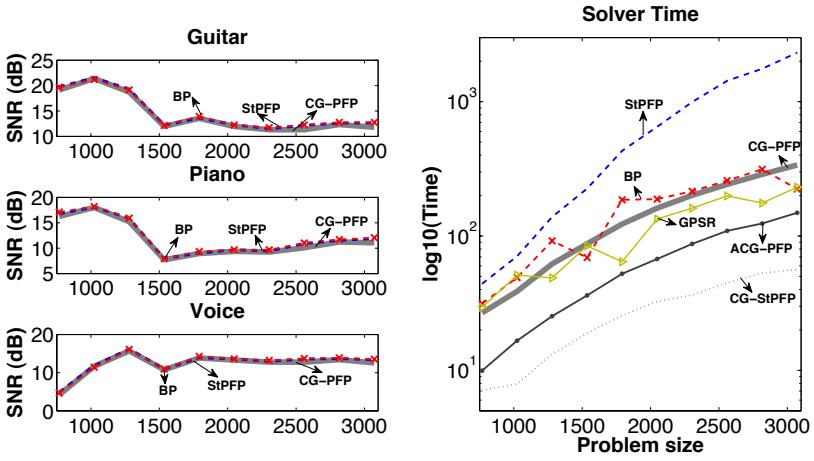


Figure 3.9: Experimental results for the modified SPARCO problem 402: the SNR values for three of the tested algorithms are shown on the left and the elapsed times to separate the sources of all algorithms on the right.

This problem allowed for variable problem sizes by modifying appropriately SPARCO problem 402. More specifically, experiments were carried out for 10 different problem sizes (from 768 samples to 3072 in 256 samples increments). The SNR of the reconstruction of

each instrument separately and the elapsed time to solve the inverse problem were then compared.

Fig. 3.9 shows the SNR results for three of the tested algorithms (BP, CG-PFP, StPFP). As can be seen, BP yields slightly better performance, however the difference in audio quality is indistinguishable. All other tested algorithms resulted in very similar SNR performance. Regarding the elapsed recovery time of the stepwise algorithms, ACG-PFP proves to be the fastest algorithm followed by GPSR. However, the stagewise CG-StPFP algorithm achieves the fastest convergence and shows its clear advantage over StPFP which is based on Cholesky factorization.

3.6 Discussion

This chapter introduced a suite of new greedy algorithms that find approximate sparse solutions to underdetermined sparse representations problems, based on the Polytope Faces Pursuit algorithm using directional updates.

The earlier implementation of the Polytope Faces Pursuit in a greedy fashion identifies the appropriate atom of the overcomplete dictionary according to the scaled correlation criterion and then updates the solution vector using the method of Cholesky factorization. This method has shown to be inefficient for large scale recovery problems, as it involves high memory requirements. The proposed method of Conjugate Gradient Polytope Faces Pursuit overcomes this problem using a conjugate gradient update at each step of the algorithm. Although the specific implementation omits the release of the retarding constraints, theoretical results showed that when ERC is satisfied both algorithms are equivalent, converging to the sparsest solution in only K iterations.

However, the dimensionality of the problem, which needs to be estimated by the conjugate gradient method increases each time the algorithm selects a new atom and therefore a full conjugate update is required at each step. The proposed Approximate Conjugate Gradient Polytope Faces Pursuit algorithm considers approximate conjugate gradient updates, reducing in that way the overall complexity and resulting in faster convergence but less accurate recovery.

Another option to reduce the complexity of the algorithm is to consider a stagewise approach. Subsequently, according to some predefined selection criteria (fixed or weak

selection criterion), the introduced Stagewise Conjugate Gradient Polytope Faces Pursuit algorithm adds several basis vectors to the already chosen subspace at each stage. The solution vector is then updated using directional updates similar to the stepwise version.

The performance of the proposed algorithms was first tested on synthetic data, using a dictionary with fixed size allowing the sparsity level and the additive noise power to vary. The simulation results showed that the proposed algorithms perform reasonably, close to the BP and GPSR performance, even in the case of approximate directional updates. Next, the performance of the proposed algorithms was compared against the stepwise Polytope Faces Pursuit algorithm based on the Cholesky factorization, in the scenario where the problem size varies while the sparsity and noise level remain fixed. The results showed that the time needed for the sparse recovery rises almost linearly with the problem size for the PFP algorithms based on directional updates, with the stagewise method being the fastest. On the other hand, the original PFP method rises faster than linear as the problem size increases.

Finally, the algorithms were tested on an image deblurring, a compressed sensing and a source separation problem from the SPARCO toolbox. The results verified once again that the Polytope Faces Pursuit algorithm converges faster when the directional updates are used and especially when combined with the stagewise approach.

To sum up, both theoretical and simulation results have proven the robustness of the proposed suite of PFP algorithms based on directional updates in recovering sparse representations. The following chapter focuses on the adaptation of the atom selection criterion so that the sparsity is enforced in blocks of atoms. In that manner, PFP is extended in order to incorporate block structures and hence exploit this additional information to achieve better sparse recovery.

Chapter 4

Recovery of structured sparse signals

Chapter 3 presented sparse recovery algorithms based on the conventional sparsity model, which assumes that the nonzero coefficients can be located anywhere in the sparse vector $\mathbf{x}$. However, recent studies [94, 95, 96, 97] have proposed extensions of the standard sparsity model, in order to improve the recoverability of signals exhibiting some underlying low dimensional structure, which is not necessarily exploited by only the sparsity assumption. The driving force behind these approaches is that if such a structure exists in the sparse domain for a class of signals, then it could be expressed by signal models that go beyond standard sparsity and yield better signal representations when this additional structure is leveraged with the conventional sparsity model. From a compressed sensing point of view, this might allow for further reduction of the number of measurements needed for perfect recovery as reported in [60].

Regarding the problem of DOA estimation and source localization, as it will be addressed in following chapters, structured sparsity can also be very beneficial. For that reason, the aim of this chapter is to investigate in detail the problem of structured sparsity and develop the appropriate algorithms that can then be directly applied to the problem of DOA estimation.

4.1 The group sparsity model

According to the general model of union of subspaces [60, 98], there exist signal representations that exhibit block structures of the nonzero entries of the sparse representation

vector. These block structures, which imply that the nonzero elements are grouped in blocks (or clusters) instead of being arbitrarily located throughout the sparse vector $\mathbf{x}$ that satisfies $\mathbf{y} = \mathbf{A}\mathbf{x}$, can be encountered in many practical scenarios. For instance, the sparse coefficients in multi-band signals [99] or harmonic signals [100] can be clustered in groups of dictionary atoms. In that special case of structured sparsity, the group sparse (or equivalently block sparse) vector $\mathbf{x}$ is treated as a concatenation of groups of coefficients with equal size d:

$$\mathbf{x} = [\underbrace{x_1 \ \dots \ x_d}_{\mathbf{x}_1^T} \ \underbrace{x_{d+1} \ \dots \ x_{2d}}_{\mathbf{x}_2^T} \ \cdots \ \underbrace{x_{N-d+1} \ \dots \ x_N}_{\mathbf{x}_P^T}]^T \tag{4.1}$$

where $\mathbf{x}_i$ denotes the i-th group and $N = Pd$. In [101] the group K-sparse vector is defined as the vector $\mathbf{x} \in \mathbb{R}^N$ that has nonzero ℓ_p norm (e.g. ℓ_2, ℓ_1 or ℓ_∞) for at most K indices out of P, namely:

$$\|\mathbf{x}\|_{p,0} = \sum_{i=1}^{P} I(\|\mathbf{x}_i\|_p > 0) \leq K \tag{4.2}$$

where $I(\cdot)$ is the indicator function. $\|\cdot\|_{p,0}$ denotes the mixed $\ell_{p,0}$ vector norm and $\|\cdot\|_p$ denotes the ℓ_p norm. It follows that the redundant dictionary $\mathbf{A}$ can also be represented as a concatenation of P block matrices:

$$\mathbf{A} = [\underbrace{\mathbf{a}_1 \ \dots \ \mathbf{a}_d}_{\mathbf{A}_1} \ \underbrace{\mathbf{a}_{d+1} \ \dots \ \mathbf{a}_{2d}}_{\mathbf{A}_2} \ \cdots \ \underbrace{\mathbf{a}_{N-d+1} \ \dots \ \mathbf{a}_N}_{\mathbf{A}_P}] \tag{4.3}$$

where $\mathbf{A}_i$ denotes the i-th column block matrix of size $M \times d$.

4.1.1 Approaches to group sparsity

In the conventional model, sparsity can be achieved by the penalization of the cost function (i.e. least squares) with the ℓ_1 norm of the coefficient vector $\mathbf{x}$. In contrast, structured sparsity may be induced by penalizing with other functions instead of the ℓ_1 norm [102]. These functions should take into account this group structure and therefore select simultaneously all nonzero entries that form a group of coefficients. Most of the existing literature mainly focuses on norms that can be written as a linear combination of norms on subsets of the group sparse vector $\mathbf{x}$. More specifically, it has been shown in

[103] that a regularization norm such as $\ell_{p,1}$, defined as:

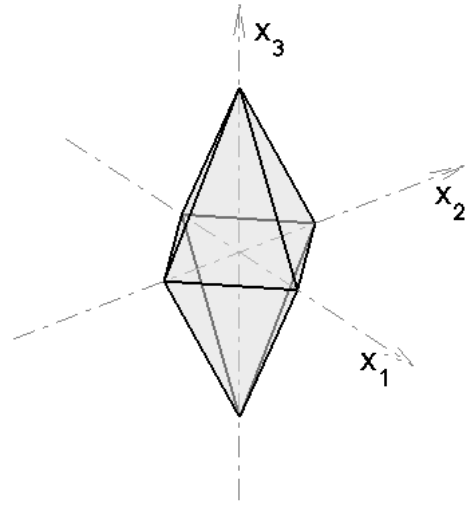
$$\|\mathbf{x}\|_{p,1} = \sum_{i=1}^{P} \|\mathbf{x}_i\|_p \tag{4.4}$$

explicitly exploits the underlying group structure and leads to improved recovery performance when the aforementioned group sparsity assumption holds. In practice, the choice of $p = 2$ seems to be the most common approach resulting in the mixed $\ell_{2,1}$ norm minimization as proposed by Eldar et al. in [101, 98]:

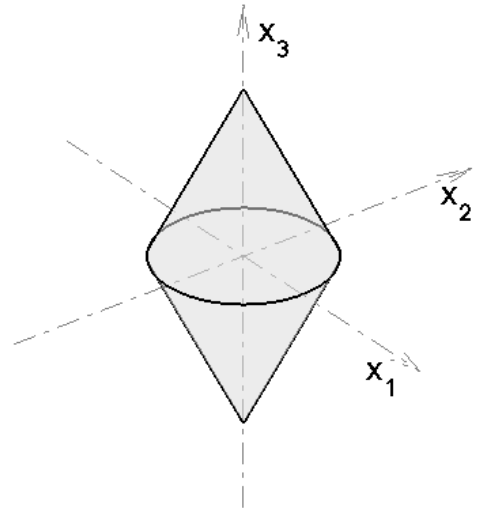
$$\min_{\mathbf{x}} \|\mathbf{x}\|_{2,1} \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{4.5}$$

The problem of (4.5) can be viewed as a natural extension of ℓ_1 minimization to the group sparsity scenario, where sparsity is enforced by minimizing the ℓ_1 norm of the $P \times 1$ vector consisting of the ℓ_2 norms of each group of coefficients. A less popular approach considers the problem of $\ell_{\infty,1}$ minimization, which simply replaces the ℓ_2 norm in the objective function with the ℓ_∞ and therefore group sparsity is induced by minimizing the ℓ_1 norm of the maximum absolute value of each group of entries in the group sparse vector $\mathbf{x}$. As discussed in [103], this method exhibits certain algorithmic benefits especially when it comes to group structures with overlapping groups of dictionary atoms. Fig. 4.1(a)-(c) show the ℓ_p balls in three dimensions of the objective functions for the conventional (ℓ_1 ball) and the group sparsity ($\ell_{2,1}$ and $\ell_{\infty,1}$ balls).

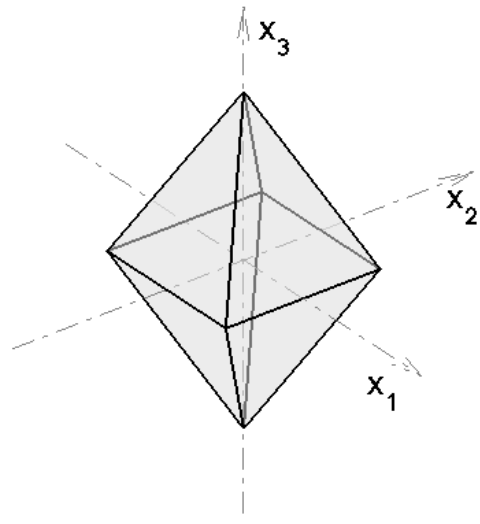
Since the objective function is convex, the problem of (4.5) is a convex optimization problem and can be solved in polynomial time. Therefore, in [104] (4.5) is cast as a second order cone program (SOCP), whereas in [105] it was shown that it can also be appropriately formulated as a semi-definite program (SDP). In both cases the solution can be efficiently retrieved using standard software packages (i.e. SeDuMi [106], SDPT3 [107]). However, although these packages can yield high precision solutions, they are usually adopted for low scale problems. As discussed in [103], when dealing with large scale problems, simpler iterative methods are preferred for their faster convergence. Although these methods usually tend to yield lower precision solutions, they are sufficient for most practical applications. For this reason, Block Coordinate Descent algorithms (i.e. group soft thresholding) and homotopy based methods have been proposed, as discussed in



(a)



(b)



(c)

Figure 4.1: Three dimensional balls of norms enforcing sparsity: (a) ℓ_1 ball, (b) $\ell_{2,1}$ ball and (c) $\ell_{\infty,1}$ ball.

[103]. Of particular interest to the motives of this research is the work in [108] and [109], where the authors proposed the group Lasso and group LARS algorithms respectively, which stand as extensions of the original conventional sparsity algorithms.

On the other hand, greedy algorithms can serve as alternatives to the optimization in equation (4.5) e.g. Group Orthogonal Matching Pursuit (GOMP) [101]. As its name suggests, GOMP is the natural extension of OMP to the group sparsity case. In more detail, the algorithm at the k-th step selects the group of atoms that best matches the current residual according to the criterion:

$$i^k = \arg \max_{i \notin \mathcal{I}^{k-1}} \|\mathbf{A}_i^T \mathbf{r}^{k-1}\|_2 \tag{4.6}$$

where $\mathbf{r}^{k-1}$ is the current residual, initialized as $\mathbf{r}^0 = \mathbf{y}$. Once the index i^k has been identified, it is added to the set of selected indices $\mathcal{I}^{k-1}$ and then the algorithm updates the solution vector $\mathbf{x}^k$ with the solution to the least squares problem:

$$\min_{\mathbf{x}^k} \left\| \mathbf{y} - \sum_{i \in \mathcal{I}^k} \mathbf{A}_i \mathbf{x}_i^k \right\|_2 . \tag{4.7}$$

Next, as in OMP the new residual is estimated $\mathbf{r}^k = \mathbf{y} - \sum_{i \in \mathcal{I}^k} \mathbf{A}_i \mathbf{x}_i^k$ and if $\mathbf{r}^k = \mathbf{0}$ the algorithm terminates. It has been shown in [101], that for both mixed $\ell_{2,1}$ minimization and GOMP recovery conditions exist that unveil their performance gain over the corresponding conventional sparsity algorithms of ℓ_1 minimization and OMP, respectively.

In Chapter 3, it has been shown that PFP shares similar theoretical properties with BP. Experimental results showed that both PFP and its suboptimal version of CG-PFP in certain cases can yield performance as good as BP. However, PFP in its original version makes no assumptions about the structure of the sparse coefficients vector $\mathbf{x}$. Therefore, the question that follows naturally is whether one can build robust algorithms based on PFP that enhance its performance in the group sparsity scenario and provide faster alternatives to the existing group sparse recovery methods. In what follows, based on the GOMP algorithm and the homotopy based method of [109], the group sparse structure is in more detail investigated and eventually, the standard PFP algorithm is extended to the group sparsity case. Theoretical and empirical results demonstrate the robustness of the proposed algorithms.

4.2 Recovery of group sparse signals via Group Polytope Faces Pursuit

In this section, the Group Polytope Faces Pursuit (GPFP) is developed using two different constraints for group sparsity, namely $\ell_{2,1}$ and $\ell_{\infty,1}$. The dual linear program for each case is first derived and then the corresponding algorithm is developed. In this discussion, it is assumed that the groups of coefficients are disjoint and of equal length d. Therefore, overlapping groups are not considered.

4.2.1 Dual linear program of the $\ell_{2,1}$ group sparse recovery problem

Before the development of the Group Polytope Faces Pursuit algorithm, it is first required to derive the dual of (4.5) along with the necessary and sufficient conditions for primal-dual optimality. To do this, the Lagrangian associated with the problem (4.5) is considered. The Lagrangian of a constrained problem can be found by augmenting the objective function with a weighted sum of the constraints [2]. Therefore, for the problem of equation (4.5), it is given by:

$$\mathcal{L}(\mathbf{x}, \mathbf{c}) = \|\mathbf{x}\|_{2,1} + \mathbf{c}^T(\mathbf{y} - \mathbf{A}\mathbf{x}) \tag{4.8}$$

where $\mathbf{c}$ is the dual vector of (4.5), containing the Lagrangian multipliers associated with the equality constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$. It is then essential to find the minimum value of the Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{c})$ over $\mathbf{x}$ and for that reason the dual function $g(\mathbf{c})$ is defined as:

$$g(\mathbf{c}) = \inf_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{c}) = \inf_{\mathbf{x}} \{\|\mathbf{x}\|_{2,1} + \mathbf{c}^T(\mathbf{y} - \mathbf{A}\mathbf{x})\} = \mathbf{c}^T\mathbf{y} + \inf_{\mathbf{x}} \{\|\mathbf{x}\|_{2,1} - \mathbf{c}^T\mathbf{A}\mathbf{x}\}. \tag{4.9}$$

Further manipulation of the above equation shows that:

$$g(\mathbf{c}) = \mathbf{c}^T\mathbf{y} - \sup_{\mathbf{x}} \{\mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{2,1}\} = \mathbf{c}^T\mathbf{y} - \begin{cases} 0 & \text{if} \quad \|\mathbf{A}^T\mathbf{c}\|_{2,\infty} \leq 1 \\ \infty & \text{otherwise} \end{cases}. \tag{4.10}$$

This can be easily proven. Define $f(\mathbf{c}) = \sup_{\mathbf{x}} \{\mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{2,1}\}$, which is the conjugate function of the objective function as explained in [2]. When $\|\mathbf{c}^T\mathbf{A}\|_{2,\infty} > 1$, there exists a $\mathbf{z} \in \mathbb{R}^N$ with $\|\mathbf{z}\|_{2,1} \leq 1$ and $\mathbf{c}^T\mathbf{A}\mathbf{z} > 1$, which follows from the definition of the

dual norm. Taking $\mathbf{x} = t\mathbf{z}$ and letting $t \to \infty$, it becomes clear that $\mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{2,1} = t(\mathbf{c}^T\mathbf{A}\mathbf{z} - \|\mathbf{z}\|_{2,1}) \to \infty$. On the other hand, when $\|\mathbf{c}^T\mathbf{A}\|_{2,\infty} \leq 1$ it holds that $\mathbf{c}^T\mathbf{A}\mathbf{x} \leq \|\mathbf{x}\|_{2,1}\|\mathbf{A}^T\mathbf{c}\|_{2,\infty}$ for any $\mathbf{x}$. This implies that $\mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{2,1} \leq \|\mathbf{x}\|_{2,1}(\|\mathbf{A}^T\mathbf{c}\|_{2,\infty} - 1) \leq 0$. Consequently, the vector that maximizes $\mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{2,1}$ is $\mathbf{x} = \mathbf{0}$. Therefore, as shown in [2] in the more general case, for the conjugate function $f(\mathbf{c})$, it holds:

$$f(\mathbf{c}) = \begin{cases} 0 & \text{if} \quad \|\mathbf{A}^T\mathbf{c}\|_{2,\infty} \leq 1 \\ \infty & \text{otherwise} \end{cases}. \tag{4.11}$$

It is obvious that the Lagrangian dual vector $\mathbf{c}$ is feasible if $\|\mathbf{A}^T\mathbf{c}\|_{2,\infty} \leq 1$. In any other case the dual function $g(\mathbf{c})$ is not bounded from above and $\mathbf{c}$ is infeasible. In the feasible case, the dual function $g(\mathbf{c}) \leq \mathbf{c}^{*T}\mathbf{y}$ gives a lower bound on the optimum solution $\mathbf{c}^*$. In order to obtain the dual program of (4.5), the best lower bound is required. This leads to the optimization problem:

$$\max_{\mathbf{c}} \mathbf{y}^T\mathbf{c} \quad \text{such that} \quad \|\mathbf{A}^T\mathbf{c}\|_{2,\infty} \leq 1. \tag{4.12}$$

Notice that the constraint in (4.12), considers the mixed $\ell_{2,\infty}$ norm, which is dual to the $\ell_{2,1}$ norm of the objective function in (4.5). Actually, the ℓ_1 and ℓ_∞ norms are dual to each other, while the ℓ_2 norm is dual to itself. It follows naturally that the mixed $\ell_{2,1}$ and $\ell_{2,\infty}$ will be dual to each other.

Suppose now that $\mathbf{x}^*$ and $\mathbf{c}^*$ is the primal-dual optimal pair of problems (4.5) and (4.12) respectively. As it was shown, the Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{c})$ takes its minimum value for the pair $\mathbf{x}^*$ and $\mathbf{c}^*$ and its gradient should vanish at $\mathbf{x}^*$. It follows that the KKT optimality conditions require that:

$$\mathbf{y} = \mathbf{A}\mathbf{x}^* \quad \text{and} \quad \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \mathbf{c}^*) = \mathbf{0}. \tag{4.13}$$

After differentiating $\mathcal{L}(\mathbf{x}^*, \mathbf{c}^*)$ with respect to $\mathbf{x}$ the second condition becomes:

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \mathbf{c}^*) = \nabla\|\mathbf{x}^*\|_{2,1} - \mathbf{A}^T\mathbf{c}^* = \mathbf{0} \tag{4.14}$$

where $\nabla\|\mathbf{x}^*\|_{2,1}$ denotes the subgradient of $\|\mathbf{x}^*\|_{2,1}$. In general, ∇f is the subgradient of

a convex function $f(\mathbf{x})$ at $\mathbf{x}$ for any $\mathbf{z}$ if $f(\mathbf{z}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{z}-\mathbf{x})$. When the function $f(\mathbf{x})$ is convex and differentiable then its gradient is also a subgradient. However, the subgradient can also exist even if $f(\mathbf{x})$ is not differentiable at some $\mathbf{x}$, as illustrated in Fig. 4.2. As can be seen, around x_1, the convex function $f(z)$ is smooth and therefore differentiable. As a result, only one subgradient exists at this point. On the other hand, around x_2, $f(z)$ is non-smooth and several subgradients exist. Moreover, it can be noted that each subgradient specifies an affine function which is tangent to the function $f(z)$ at any given point.

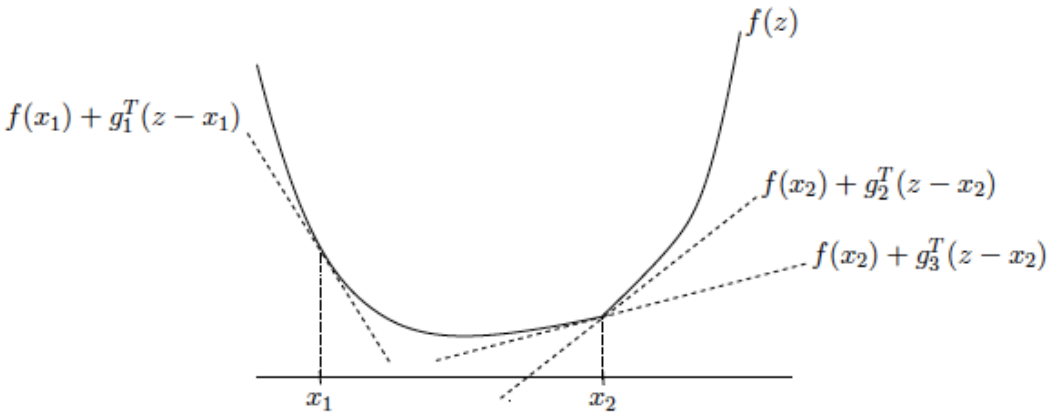


Figure 4.2: At x_1, the convex function f is differentiable, and g_1 (which is the derivative of f at x_1) is the unique subgradient at x_1. At the point x_2, f is not differentiable. At this point, f has many subgradients: two subgradients, g_2 and g_3, are shown [2].

Therefore, if $f(\mathbf{x})$ is convex and differentiable, then its gradient is its only subgradient, whereas if $f(\mathbf{x})$ is not differentiable, there might exist more than one subgradient. In the latter case, which in practice arises when dealing with non-smooth functions, it is more convenient to work with subdifferentials. A function $f(\mathbf{x})$ is subdifferentiable at $\mathbf{x}$ if there exists at least one subgradient. This implies that a function $f(\mathbf{x})$ can be subdifferentialable at $\mathbf{x}$ even if it is not differentialable. The subdifferential ∂f therefore generalizes the concept of subgradients and it is defined as the set of all subgradients of f at $\mathbf{x}$. The subdifferential of f at $\mathbf{x}$ can then be more formally defined as the set:

$$\partial f(\mathbf{x}) = \left\{ \mathbf{g} \in \mathbb{R}^N \mid f(\mathbf{x}) + \mathbf{g}^T(\mathbf{z}-\mathbf{x}) \leq f(\mathbf{z}), \quad \forall \mathbf{x} \in \mathbb{R}^N \right\}. \tag{4.15}$$

Note that if f is differentiable at $\mathbf{x}$, then the set reduces to a unit set (or singleton).

Considering that the function of interest, namely $\|\mathbf{x}\|_{2,1}$, is not differentiable at $\mathbf{x} = \mathbf{0}$

it is more handy to work with subdifferenitals instead of subgradients. More specifically, it can easily be shown that the gradient of the i-th group $\nabla\|\mathbf{x}_i\|_{2,1}$ is given by the expression $\nabla\|\mathbf{x}_i\|_{2,1} = \mathbf{x}_i/\|\mathbf{x}_i\|_2$ when $\|\mathbf{x}_i\|_2 > 0$. It follows that when $\|\mathbf{x}_i\|_2 > 0$, the subdifferential is given by $\partial\|\mathbf{x}_i\|_{2,1} = \mathbf{x}_i/\|\mathbf{x}_i\|_2$. However, for the zero block-elements of $\mathbf{x}$ the gradient is not defined, but $\partial\|\mathbf{x}_i\|_{2,1}$ coincides with the set of unit ℓ_2 norm vectors $\mathcal{B}^d_{\ell_2} = \{\mathbf{u} \in \mathbb{R}^d \mid \|\mathbf{u}\|_2 \leq 1\}$ [110]. Therefore, for each $i = 1, \ldots, P$, it holds:

$$\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{2,1} = \begin{cases} \{\mathbf{u} \in \mathbb{R}^d \mid \|\mathbf{u}\|_2 = 1 : \mathbf{u} = \mathbf{x}_i/\|\mathbf{x}_i\|_2\} & \text{if} \quad \|\mathbf{x}_i\|_2 > 0 \\ \\ \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\|_2 \leq 1\} & \text{otherwise} \end{cases}. \tag{4.16}$$

The second KKT condition of (4.13) can then be rewritten as $\partial_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \mathbf{c}^*) \in \mathbf{0}$ and equation (4.14) becomes:

$$\partial_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \mathbf{c}^*) = \partial\|\mathbf{x}^*\|_{2,1} - \mathbf{A}^T\mathbf{c}^* \in \mathbf{0}. \tag{4.17}$$

Substituting equation (4.16) to (4.17), it yields:

$$\mathbf{A}_i^T\mathbf{c}^* \in \partial_{\mathbf{x}_i}\|\mathbf{x}_i^*\|_{2,1} \tag{4.18}$$

and therefore $\|\mathbf{A}^T\mathbf{c}^*\|_{2,\infty} \leq 1$. It has been shown that for the optimal $\mathbf{x}^*$, there exists a corresponding optimal $\mathbf{c}^*$. According to the KKT conditions for primal-dual feasibility, the following lemma holds.

Lemma 2. *Suppose that the primal problem of equation* (4.5) *is feasible. Then the pair* $\mathbf{x}^*$, $\mathbf{c}^*$ *is an optimum point for both primal and dual linear programs if and only if the following conditions hold:*

$$\mathbf{A}\mathbf{x}^* = \mathbf{y} \tag{4.19a}$$

$$\|\mathbf{A}^T\mathbf{c}^*\|_{2,\infty} \leq 1 \tag{4.19b}$$

$$\mathbf{y}^T\mathbf{c}^* = \|\mathbf{x}^*\|_{2,1}. \tag{4.19c}$$

Proof. This follows immediately from KKT conditions. Condition (4.19c) can be derived

from (4.18). Define $\mathbf{v}^* = [\partial_{\mathbf{x}_1}\|\mathbf{x}_1^*\|_{2,1}\ \partial_{\mathbf{x}_2}\|\mathbf{x}_2^*\|_{2,1}\ \ldots\ \partial_{\mathbf{x}_P}\|\mathbf{x}_P^*\|_{2,1}]^T$. Then, from (4.18) it follows that $\mathbf{A}^T\mathbf{c}^* = \mathbf{v}^* \Rightarrow (\mathbf{x}^*)^T\mathbf{A}^T\mathbf{c}^* = (\mathbf{x}^*)^T\mathbf{v}^* \Rightarrow \mathbf{y}^T\mathbf{c}^* = \|\mathbf{x}^*\|_{2,1}$. □

Lemma 2 provides the necessary and sufficient conditions for primal-dual optimality. More specifically, conditions (4.19a) and (4.19b) concern the primal and dual feasibility. Condition (4.19c) provides the complementary slackness, which implies zero duality gap and therefore strong duality holds between the primal and the dual problems. According to [2], strong duality between the primal and its dual problem suggests that the primal and dual solutions are equivalent.

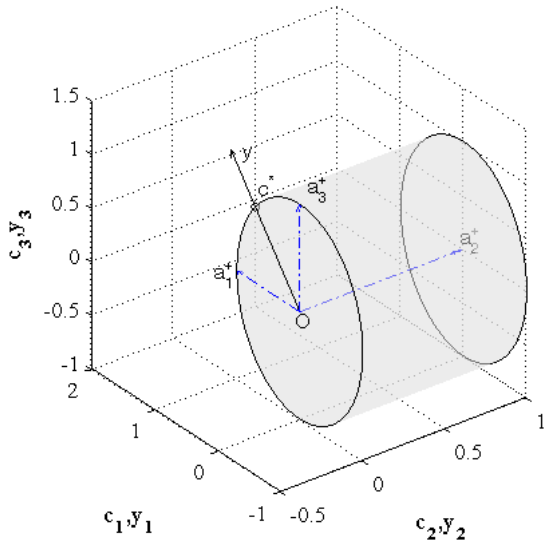


Figure 4.3: Geometry of the dual of the mixed $\ell_{2,1}$ minimization problem in three dimensions. The dictionary atoms $\mathbf{a}_1$ and $\mathbf{a}_3$ form a group, while $\mathbf{a}_2$ is treated as in the conventional sparsity model assuming also nonnegativity for its corresponding coefficient. The optimum dual vector $\mathbf{c}^*$ lies on the circumference of the circular surface defined by $\mathbf{a}_1$ and $\mathbf{a}_3$ (section of cylinder at $c_2 = 0$).

Fig. 4.3 illustrates the geometry of the dual linear program for the $\ell_{2,1}$ minimization problem in three dimensions, where it is supposed that the dictionary atoms $\mathbf{a}_1$ and $\mathbf{a}_3$ form a group, while the atom $\mathbf{a}_2$ is treated as in the case of conventional sparsity with the additional nonnegativity constraint, since only positive values along the direction of $\mathbf{a}_2$ are considered. As can be seen, the geometrical entity associated with the dual linear program is a cylinder and for the given measurements vector $\mathbf{y}$ the optimum dual vector $\mathbf{c}^*$ lies on the unit circle defined by the inequality $\|[\mathbf{a}_1\ \mathbf{a}_3]^T\mathbf{c}\|_2 = 1$. Furthermore, in the specific case that the measurements vector $\mathbf{y}$ bisects the angle between the orthonormal

grouped atoms, the optimum $\mathbf{c}^*$ falls onto the intersection of $\mathbf{y}$ with the unit circle.

As in the case of conventional sparsity, the GPFP algorithm is based on the geometry associated with the dual linear program and searches the optimum vector $\mathbf{c}^*$ using a path following approach. In the following section the proposed algorithm is presented.

4.2.2 The proposed algorithm with $\ell_{2,1}$ objective

In this section, the proposed algorithm for recovery of block sparse signals based on $\ell_{2,1}$ minimization is derived. GPFP is an iterative greedy algorithm that builds the solution vector in a similar way to the GOMP algorithm. As in the case of the conventional sparsity, the algorithm starts from the origin O and follows the path towards the observations vector $\mathbf{y}$. Therefore, at the first iteration the algorithm will move along the path $\mathbf{h} = \lambda\mathbf{y}$ and the first curved face of the corresponding polytope (or more precisely intersection of cylinders) will be encountered at:

$$\lambda_1 = \min_{i \notin \mathcal{I}^0} \left\{ \|\mathbf{A}_i^T \mathbf{h}\|_2 = 1 \right\} = \min_{i \notin \mathcal{I}^0} \left\{ \|\mathbf{A}_i^T (\lambda \mathbf{y})\|_2 = 1 \right\}. \tag{4.20}$$

Further consideration shows that equation (4.20) becomes:

$$\lambda_1 = \min_{i \notin \mathcal{I}^0} \left\{ \|\mathbf{A}_i^T (\mathbf{y})\|_2 = 1/\lambda \right\} = 1/\max_{i \notin \mathcal{I}^0} \|\mathbf{A}_i^T \mathbf{y}\|_2. \tag{4.21}$$

Note that at the first iteration GPFP incorporates the ℓ_2 norm of the residual correlations of each group and therefore this step is identical to the GOMP algorithm. It turns out that both algorithms will select the same group of dictionary atoms at the first iteration. After adding the selected group into the active set, the algorithm will update the solution vector $\mathbf{x}^1$ and the corresponding vector $\mathbf{c}^1$. As in the case of the standard sparsity algorithm, the estimation of $\mathbf{x}^1$ requires a single least squares inversion, whereas for the update of $\mathbf{c}^1$ the algorithm needs to solve $\|(\mathbf{A}^1)^T \mathbf{c}^1\|_2 = 1$, so that condition (4.19b) is satisfied. This is a quadratic system and can be solved, using trust-region methods [111, 112].

As opposed to GOMP, at the next iteration instead of projecting along the current residual $\mathbf{r}^1$ starting again from the origin, the algorithm projects along the first curved surface, defined by the selected group of atoms at the first stage, towards $\mathbf{r}^1$. In that

way, although its path is specified by the direction of the residual it is constrained to be within the polar polytope so that it always satisfies condition (4.19b) of Lemma 2. Thus, the next group of dictionary atoms $\mathbf{A}_i$ will be encountered at:

$$\lambda_2 = \min_{i \notin \mathcal{I}^1} \left\{ \|\mathbf{A}_i^T \mathbf{h}\|_2 = 1 \right\} = \min_{i \notin \mathcal{I}^1} \left\{ \|\mathbf{A}_i^T (\mathbf{c}^1 + \lambda \mathbf{r}^1)\|_2 = 1 \right\} \tag{4.22}$$

where $\mathbf{r}^1$ is the current residual. Hence, more generally the algorithm at the k-th step needs to solve the above quadratic equation of (4.22) for each group of atoms that has not already been selected in previous iterations and find the minimum λ among all potential groups. Next, the algorithm adds the selected group of atoms to the active set and updates the solution vector $\mathbf{x}^k$, the residual $\mathbf{r}^k$ and the corresponding $\mathbf{c}^k$. The algorithm iterates till the stopping criteria are met. The resulting algorithm of GPFP is given in Algorithm 5.

Algorithm 5 Group Polytope Faces Pursuit (GPFP - $\ell_{2,1}$) with quadratic updates

Input: $\mathbf{A} = [\mathbf{a}_i]$, $\mathbf{y}$
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize:
$k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \mathbf{x}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \mathbf{a}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next group}
  $k \leftarrow k + 1$
  Find group of atoms:
  $i^k \leftarrow \arg\min_{i \notin \mathcal{I}^{k-1}} \{\|\mathbf{A}_i^T (\mathbf{c}^{k-1} + \lambda \mathbf{r}^{k-1})\|_2 = 1\}$
  Add constraints:
  $\mathbf{A}^k \leftarrow [\mathbf{A}_i^{k-1},\ \mathbf{A}_i^k],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\mathbf{x}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{y},\ \hat{\mathbf{y}}^k \leftarrow \mathbf{A}^k \mathbf{x}^k,\ \mathbf{r}^k \leftarrow \mathbf{y} - \hat{\mathbf{y}}^k$
  $\mathbf{c}^k \leftarrow \{\mathbf{c}^k \mid \|(\mathbf{A}^k)^T \mathbf{c}^k\|_2 = 1\}$
end while
Output: $\mathbf{x}^* \leftarrow \mathbf{x}^k$

Thoroughly inspecting Algorithm 5, it can be noted that GPFP at each iteration will satisfy the primal and dual feasibility conditions ((4.19a), (4.19b)) of Lemma 2, but it is not guaranteed that the complementary slackness condition of equation (4.19c) will be satisfied. In other words, Algorithm 5 does not guarantee convergence to the optimum $\mathbf{c}^*$, when the conditions of Lemma 2 is met. In order for the GPFP algorithm to satisfy condition (4.19c), the a priori knowledge of the unknown optimum sparse vector $\mathbf{x}^*$ should have been required, so that the algorithm can estimate the subdifferential of equation (4.16) and hence update the solution vector $\mathbf{c}^k$ by solving the following least squares

problem:

$$(\mathbf{A}^k)^T\mathbf{c}^k = \mathbf{v}^k \tag{4.23}$$

where $\mathbf{v}^k$ is a vector containing the concatenation of subdifferentials of the selected k groups of atoms, defined as:

$$\mathbf{v}_i^k = \partial_{\mathbf{x}_i}\|\mathbf{x}_i^*\|_{2,1} = \mathbf{x}_i^*/\|\mathbf{x}_i^*\|_2 \tag{4.24}$$

for $i = 1, \ldots, k$. The same also applies for the step of identifying the new group of atoms to enter the active set. Ideally, the atom selection criterion should be posed as $\min\{\mathbf{A}_i^T(\mathbf{c}^{k-1} + \lambda\mathbf{r}^{k-1}) = \mathbf{v}_i^k\}$. For that reason, Algorithm 5 should be viewed as a suboptimal greedy method to the problem of equation 4.5. Nevertheless, as it will also be addressed in Section 4.3, stronger conditions than the ones of Lemma 2 exist that can give some more insight into the theoretical properties of the proposed method.

Despite that, the above observation can lead to a valuable realization that can in turn result in decreasing the computational cost of Algorithm 5. First of all, notice that the most computationally expensive operations are the quadratic problems encountered at the group selection step and the update of vector $\mathbf{c}$. More specifically, at each iteration GPFP needs to solve $P - k$ quadratic problems required for the group selection criterion involving one unknown parameter at a time and an additional quadratic problem to compute $\mathbf{c}^k \in \mathbb{R}^M$. Having mentioned that this approach is already suboptimal, one can attempt to find an approximate estimate of the subdifferential vector $\partial_{\mathbf{x}_i}\|\mathbf{x}_i^*\|_{2,1}$ and therefore reduce the specific quadratic problems to typical linear least squares problems. To do this, an obvious choice is to consider the normalized correlation $\mathbf{A}_i^T\mathbf{r}^{k-1}/\|\mathbf{A}_i^T\mathbf{r}^{k-1}\|_2$ so that its ℓ_2 norm is unit as required by the algorithm. In that case, equation (4.22) becomes:

$$\begin{aligned}
\lambda_k &= \min_{i\notin\mathcal{I}^{k-1}} \left\{\mathbf{A}_i^T\mathbf{h}^k = \mathbf{A}_i^T\mathbf{r}^{k-1}/\|\mathbf{A}_i^T\mathbf{r}^{k-1}\|_2\right\} \\
&= \min_{i\notin\mathcal{I}^{k-1}} \left\{\mathbf{A}_i^T(\mathbf{c}^{k-1} + \lambda\mathbf{r}^{k-1}) = \frac{\mathbf{A}_i^T\mathbf{r}^{k-1}}{\|\mathbf{A}_i^T\mathbf{r}^{k-1}\|_2}\right\} \\
&= \min_{i\notin\mathcal{I}^{k-1}} \left\{(\mathbf{r}^{k-1})^T\mathbf{A}_i\mathbf{A}_i^T(\mathbf{c}^{k-1} + \lambda\mathbf{r}^{k-1}) = \|\mathbf{A}_i^T\mathbf{r}^{k-1}\|_2\right\}
\end{aligned} \tag{4.25}$$

and accordingly the selection criterion at the k-th iteration will be:

$$\alpha_k = \arg \max_{i \notin \mathcal{I}^{k-1}} \frac{\|\mathbf{A}_i^T \mathbf{r}^{k-1}\|_2^2}{\|\mathbf{A}_i^T \mathbf{r}^{k-1}\|_2 - (\mathbf{r}^{k-1})^T \mathbf{A}_i \mathbf{A}_i^T \mathbf{c}^{k-1}} \tag{4.26}$$

where the scalar α_k is defined as the inverse of λ_k. Note that when the block size is $d = 1$ equation (4.26) reduces to the maximum scaled correlation of equation (3.7). In accordance with Algorithm 5, after the new group of atoms has been selected the algorithm will proceed in updating the solution vectors $\mathbf{x}^k$, $\mathbf{r}^k$ and $\mathbf{c}^k$. Thus, the vector $\mathbf{v}^k$ can be constructed by concatenating the vectors $\mathbf{A}_{i^k}^T \mathbf{r}^{k-1} / \|\mathbf{A}_{i^k}^T \mathbf{r}^{k-1}\|_2$ containing normalized correlations for any group indexed by i^k that is added in the active set at each iteration of the algorithm, namely:

$$\mathbf{v}^k = \left[\frac{\mathbf{A}_{i^1}^T \mathbf{r}^0}{\|\mathbf{A}_{i^1}^T \mathbf{r}^0\|_2} \; \frac{\mathbf{A}_{i^2}^T \mathbf{r}^1}{\|\mathbf{A}_{i^2}^T \mathbf{r}^1\|_2} \; \cdots \; \frac{\mathbf{A}_{i^k}^T \mathbf{r}^{k-1}}{\|\mathbf{A}_{i^k}^T \mathbf{r}^{k-1}\|_2} \right]^T. \tag{4.27}$$

One can then substitute $\mathbf{v}^k$ to equation (4.23) and solve the resulting system of linear equations for $\mathbf{c}^k$. The modified GPFP algorithm is illustrated in Algorithm 6.

Algorithm 6 Group Polytope Faces Pursuit (GPFP - $\ell_{2,1}$)

Input: $\mathbf{A} = [\mathbf{a}_i]$, $\mathbf{y}$
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize:
  $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \mathbf{x}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y},\ \mathbf{v}^k \leftarrow \emptyset$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \mathbf{a}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next group}
  $k \leftarrow k + 1$
  Find group of atoms:
  $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \|\mathbf{A}_i^T \mathbf{r}^{k-1}\|_2^2 / (\|\mathbf{A}_i^T \mathbf{r}^{k-1}\|_2 - (\mathbf{r}^{k-1})^T \mathbf{A}_i \mathbf{A}_i^T \mathbf{c}^{k-1})$
  Add constraints:
  $\mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{A}_i^k],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\mathbf{x}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{y},\ \hat{\mathbf{y}}^k \leftarrow \mathbf{A}^k \mathbf{x}^k,\ \mathbf{r}^k \leftarrow \mathbf{y} - \hat{\mathbf{y}}^k$
  $\mathbf{v}^k \leftarrow [\mathbf{v}^{k-1},\ \mathbf{A}_{i^k}^T \mathbf{r}^{k-1} / \|\mathbf{A}_{i^k}^T \mathbf{r}^{k-1}\|_2],\ \mathbf{c}^k \leftarrow [(\mathbf{A}^k)^\dagger]^T \mathbf{v}^k$
end while
Output: $\mathbf{x}^* \leftarrow \mathbf{x}^k$

For the implementation of Algorithm 6, Cholesky factorization has been used for updating both $\mathbf{x}^k$ and $\mathbf{c}^k$ vectors. As has already been discussed in the previous chapter for the conventional sparsity PFP algorithm, directional updates could be used (e.g. the method of conjugate gradient) when dealing with large scale systems.

4.2.3 Dual linear program of the $\ell_{\infty,1}$ group sparse recovery problem

Another possibility for recovery of group sparse signals arises by considering minimization of the $\ell_{\infty,1}$ instead of the $\ell_{2,1}$. As already mentioned and discussed in more detail in [103], in certain cases $\ell_{\infty,1}$ minimization might exhibit better group sparse recovery performance than $\ell_{2,1}$ minimization. Since the appropriate design of the norms that induce group sparsity structures seems to be an open problem with ongoing research, in what follows, based on the analysis of the previous section, the GPFP algorithm is modified by replacing the objective function with the $\ell_{\infty,1}$ norm.

In the case of $\ell_{\infty,1}$ minimization, the primal group sparse recovery problem can be formulated as:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\infty,1} \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{4.28}$$

Following a similar analysis as in Section 4.2.1, the dual function is defined as:

$$g(\mathbf{c}) = \inf_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{c}) = \inf_{\mathbf{x}} \left\{ \|\mathbf{x}\|_{\infty,1} + \mathbf{c}^T(\mathbf{y} - \mathbf{A}\mathbf{x}) \right\} = \mathbf{c}^T\mathbf{y} + \inf_{\mathbf{x}} \left\{ \|\mathbf{x}\|_{\infty,1} - \mathbf{c}^T\mathbf{A}\mathbf{x} \right\} \tag{4.29}$$

for which it holds:

$$g(\mathbf{c}) = \mathbf{c}^T\mathbf{y} - \sup_{\mathbf{x}} \left\{ \mathbf{c}^T\mathbf{A}\mathbf{x} - \|\mathbf{x}\|_{\infty,1} \right\} = \mathbf{c}^T\mathbf{y} - \begin{cases} 0 & \text{if} \quad \|\mathbf{A}^T\mathbf{c}\|_{1,\infty} \leq 1 \\ \infty & \text{otherwise} \end{cases}. \tag{4.30}$$

In that way, using equation (4.30) it can be easily shown that the dual linear program of (4.28) is given by:

$$\max_{\mathbf{c}} \mathbf{y}^T\mathbf{c} \quad \text{such that} \quad \|\mathbf{A}^T\mathbf{c}\|_{1,\infty} \leq 1. \tag{4.31}$$

Comparing now (4.31) with the dual for the $\ell_{2,1}$ norm minimization case given in (4.12), it can be noted that the main difference emerges when it comes to the constraint. This implies that the geometry associated with the dual of (4.28) will also differ from what presented in Section 4.2.1 regarding the $\ell_{2,1}$ objective.

It is noted that the KKT optimality conditions this time require that $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $\partial\|\mathbf{x}\|_{\infty,1} - \mathbf{A}^T\mathbf{c} \in \mathbf{0}$, where the subdifferential $\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{\infty,1}$ denotes the set of all subgra-

dients and it is given by:

$$\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{\infty,1} = \begin{cases} \{\mathbf{u}\in\mathbb{R}^d : \|\mathbf{u}\|_1 = 1 : \begin{cases} u_l x_l \geq 0 & \text{if} \quad |x_l| = \|\mathbf{x}_i\|_\infty \\ u_l = 0 & \text{otherwise} \end{cases}\} & , \|\mathbf{x}_i\|_\infty \neq 0 \\ \{\mathbf{u}\in\mathbb{R}^d : \|\mathbf{u}\|_1 \leq 1\} & , \|\mathbf{x}_i\|_\infty = 0 \end{cases} \tag{4.32}$$

where the index l refers to the l-th entry of the vectors $\mathbf{x}_i \in \mathbb{R}^d$ and $\mathbf{u} \in \mathbb{R}^d$.

Therefore, for the optimal $\mathbf{x}^*$ there exists a corresponding optimal $\mathbf{c}^*$ and according to the KKT conditions for this primal-dual optimal pair, the following lemma can be more formally stated.

Lemma 3. *Suppose that the primal problem of equation* (4.28) *is feasible. Then the pair* $\mathbf{x}^*$,$\mathbf{c}^*$ *is an optimum point for both primal and dual linear programs if and only if the following conditions hold:*

$$\mathbf{A}\mathbf{x}^* = \mathbf{y} \tag{4.33a}$$

$$\|\mathbf{A}^T\mathbf{c}^*\|_{1,\infty} \leq 1 \tag{4.33b}$$

$$\mathbf{y}^T\mathbf{c}^* = \|\mathbf{x}\|_{\infty,1}. \tag{4.33c}$$

Proof. This follows immediately from KKT conditions. □

Consequently, as in the case of $\ell_{2,1}$ minimization (Lemma 2), the conditions (4.33a) and (4.33b) impose the primal and dual feasibility respectively, while condition (4.33c) expresses the complementary slackness. These conditions are sufficient and necessary for the primal-dual optimality.

For the problem of Fig. 4.3, if the $\ell_{2,1}$ norm objective function is replaced by the $\ell_{\infty,1}$ norm the resulting polytope is the one depicted in Fig. 4.4. In this case, the optimum vector $\mathbf{c}^*$ lies on the intersection of the measurements vector $\mathbf{y}$ with the ℓ_1 ball defined by the orthonormal vectors $\mathbf{a}_1$ and $\mathbf{a}_3$, which ensures that the optimality condition $\|[\mathbf{a}_1\ \mathbf{a}_3]^T\mathbf{c}\|_1 = 1$ holds.

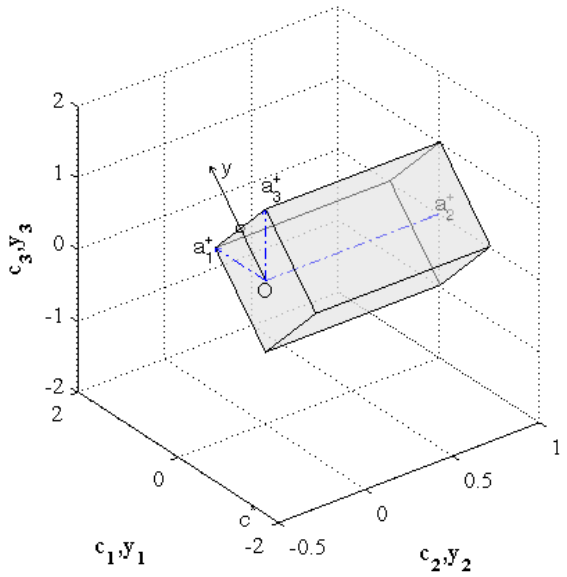


Figure 4.4: Geometry of the dual of the mixed $\ell_{\infty,1}$ minimization problem in three dimensions. The dictionary atoms $\mathbf{a}_1$ and $\mathbf{a}_3$ form a group, while $\mathbf{a}_2$ is treated as in the conventional sparsity model assuming also nonnegativity for its corresponding coefficient. The optimum dual vector $\mathbf{c}^*$ lies on one of the four faces of the surface defined by $\mathbf{a}_1$ and $\mathbf{a}_3$ (section of the polytope at $c_2 = 0$).

4.2.4 The proposed algorithm with $\ell_{\infty,1}$ objective

Once more, the geometry of the dual problem of equation (4.31) can be used to develop a GPFP algorithm for the $\ell_{\infty,1}$ norm case in a similar manner to the $\ell_{2,1}$ minimization algorithm. In the same way, using inequality (4.33b) a path following algorithm can be developed that at the k-th iteration will select the group of atoms $\mathbf{A}_i$ for any $i \notin \mathcal{I}^{k-1}$ for which:

$$\lambda_k = \min_{i \notin \mathcal{I}^{k-1}} \left\{ \|\mathbf{A}_i^T(\mathbf{c}^{k-1} + \lambda \mathbf{r}^{k-1})\|_1 = 1 \right\}. \tag{4.34}$$

On that account, after adding the selected group of atoms $\mathbf{A}_{i^k}$ to the current active set $\mathcal{I}^{k-1}$, the algorithm will move to updating vectors $\mathbf{x}^k$, $\mathbf{r}^k$ and $\mathbf{c}^k$. For the latter, it is required that $\|(\mathbf{A}^k)^T\mathbf{c}^k\|_1 = 1$. In that manner, the algorithm will iterate until some stopping threshold (e.g. ℓ_2 norm of the cost function) has been exceeded.

This implementation exhibits evident weaknesses. Indeed, from a theoretical viewpoint, the fact that the proposed algorithm cannot assure that condition 4.33c is always satisfied is one shortcoming that relegates this method to suboptimal. On the other hand, from a practical perspective, the algorithm requires to solve nonlinear least squares prob-

lems for the update of $\mathbf{c}$ and the identification of the next group of atoms to be added in the active set. These operations can add significantly to the computational complexity of the algorithm and when dealing with large scale systems they can even be unaffordable.

Therefore, as described in Section 4.2.2 for the $\ell_{2,1}$ minimization problem, one could attempt to approximate the unknown subdifferential of equation (4.32) by considering the subdifferential of the residual correlations $\mathbf{A}_i^T\mathbf{r}^{k-1}$ for each group. It follows that at the k-th iteration, the GPFP algorithm will require that:

$$\lambda_k = \min_{i \notin \mathcal{I}^{k-1}} \left\{ \mathbf{A}_i^T(\mathbf{c}^{k-1} + \lambda \mathbf{r}^{k-1}) = \mathbf{v}_i^k \right\} \tag{4.35}$$

where $\mathbf{v}_i^k \in \mathbb{R}^d$ such that $\|\mathbf{v}_i^k\|_1 = 1$ and:

$$\mathbf{v}_i^k = \begin{cases} v_l(\mathbf{A}_i^T\mathbf{r}^{k-1})_l \geq 0 & \text{if } |(\mathbf{A}_i^T\mathbf{r}^{k-1})_l| = \|\mathbf{A}_i^T\mathbf{r}^{k-1}\|_\infty \\ v_l = 0 & \text{else} \end{cases}. \tag{4.36}$$

Hence, the $\ell_{\infty,1}$ group selection criterion will be:

$$\alpha_k = \arg\max_{i \notin \mathcal{I}^{k-1}} \frac{(\mathbf{v}_i^k)^T\mathbf{A}_i^T\mathbf{r}^{k-1}}{\|\mathbf{v}_i^k\|_2^2 - (\mathbf{v}_i^k)^T\mathbf{A}_i^T\mathbf{c}^{k-1}}. \tag{4.37}$$

Accordingly, the update for vector $\mathbf{c}^k$ can be estimated by solving the least squares problem of equation (4.23), where $\mathbf{v}^k$ is now given by:

$$\mathbf{v}^k = [\mathbf{v}_{i^1}^1 \ \mathbf{v}_{i^2}^2 \ \ldots \ \mathbf{v}_{i^k}^k]^T \tag{4.38}$$

where i^k is the index of the group of atoms selected by the algorithm at the k-th iteration. The full algorithm is shown in Algorithm 7.

4.3 Theoretical properties

This section attempts to give some more insight into the theoretical properties of the proposed GPFP algorithms. First of all, the sufficient and necessary conditions for uniqueness of the optimum primal-dual pair $\mathbf{x}^*$ and $\mathbf{c}^*$ are provided. Next, based on the fact that the proposed greedy algorithms, as addressed in Section 4.2, cannot guarantee convergence to the optimum solution pair, it is shown that stronger conditions

Algorithm 7 Group Polytope Faces Pursuit (GPFP - $\ell_{\infty,1}$)

Input: $\mathbf{A} = [\mathbf{a}_i]$, $\mathbf{y}$
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize:
$k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{c}^k \leftarrow \mathbf{0},\ \mathbf{x}^k \leftarrow \emptyset,\ \hat{\mathbf{y}}^k \leftarrow \mathbf{0},\ \mathbf{r}^k \leftarrow \mathbf{y},\ \mathbf{v}^k \leftarrow \emptyset$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \mathbf{a}_i^T \mathbf{r}^{k-1} > \theta_{\min}$ **do** {Find next face}
  $k \leftarrow k + 1$
  Find group of atoms:
  $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \mathbf{v}_i^T \mathbf{A}_i^T \mathbf{r}^{k-1} / (\|\mathbf{v}_i\|_2^2 - \mathbf{v}_i^T \mathbf{A}_i^T \mathbf{c}^{k-1})$
  Add constraints:
  $\mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{A}_i^k],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\mathbf{x}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{y},\ \hat{\mathbf{y}}^k \leftarrow \mathbf{A}^k \mathbf{x}^k,\ \mathbf{r}^k \leftarrow \mathbf{y} - \hat{\mathbf{y}}^k$
  $\mathbf{v}^k \leftarrow [\mathbf{v}^{k-1},\ \mathbf{v}_{i^k}^k],\ \mathbf{c}^k \leftarrow [(\mathbf{A}^k)^\dagger]^T \mathbf{v}^k$
end while
Output: $\mathbf{x}^* \leftarrow \mathbf{x}^k$

exist, which provide guarantees for convergence of the GPFP algorithms to the optimum solution vector $\mathbf{x}^*$.

4.3.1 Uniqueness & recovery conditions

Lemmas 2 and 3 provide the optimality conditions for primal-dual feasibility for the group sparse recovery problem. Nevertheless, in agreement with the theoretical results for the joint sparsity case and the MMV problem presented in [110] they do not guarantee uniqueness of the optimum solution pair $\mathbf{x}^*$ and $\mathbf{c}^*$.

Suppose now that $\mathbf{x}_0$ has $\|\mathbf{x}_0\|_{2,0} \leq K$ nonzero groups of length d. Let $\mathbf{x}_S$ be the Kd-dimensional vector containing all nonzeros entries of $\mathbf{x}_0$ and $\mathbf{A}_S$ be the $M \times Kd$ matrix such that $\mathbf{y} = \mathbf{A}_S \mathbf{x}_S$. The following theorem gives the necessary and sufficient conditions for $\mathbf{x}_0$ to be the unique optimum solution of problem (4.5).

Theorem 8. *$\mathbf{x}_0$ is the unique optimum point of the primal linear program of (4.5) if and only if a) $\mathbf{A}_S$ has full rank and b) there exists some $\mathbf{c} \in \mathbb{R}^M$ such that*

$$\mathbf{A}_j^T \mathbf{c} = \mathbf{x}_{0j} / \|\mathbf{x}_{0j}\|_2, \quad \forall j \in S \tag{4.39a}$$

$$\|\mathbf{A}_j^T \mathbf{c}\|_2 < 1, \quad \forall j \notin S. \tag{4.39b}$$

Proof. For the if argument, note that $\mathbf{x}_0$ satisfies condition (4.19a) of Lemma 2. Subsequently, conditions (4.39a) and (4.39b) ensure that $\mathbf{c}$ satisfies the dual feasibility condition of Lemma 2. Moreover, the complementary slackness condition of equation (4.19c)

suggests that:

$$\begin{aligned}
&\mathbf{y}^T\mathbf{c} = \|\mathbf{x}_0\|_{2,1} \Rightarrow \\
&\sum_{j=1}^{N/d} (\mathbf{A}_j\mathbf{x}_{0j})^T\mathbf{c} - \sum_{j=1}^{K} \|\mathbf{x}_{0j}\|_2 = 0 \Rightarrow \\
&\sum_{j=1}^{N/d} \left[\mathbf{x}_{0j}^T\mathbf{A}_j^T\mathbf{c} - \|\mathbf{x}_{0j}\|_2\right] = 0 \Rightarrow \\
&\mathbf{x}_{0j}^T\mathbf{A}_j^T\mathbf{c} - \|\mathbf{x}_{0j}\|_2 = 0 \quad \text{for} \quad j = 1, \dots N/d.
\end{aligned} \tag{4.40}$$

It is therefore straightforward to show that equation (4.40) will be satisfied. Obviously, (4.40) holds for any group of coefficients $\mathbf{x}_{0j}$ with zero entries. For the groups with nonzero entries, equation (4.39b) holds. Finally, since $\mathbf{A}_S$ is a full rank matrix then it follows that $\mathbf{x}_0$ must be unique.

For the only if argument, suppose that $\mathbf{A}_S$ is not full rank. Consider then a small nonzero vector $\mathbf{x}_S^N \neq \mathbf{0}$ in the null space of $\mathbf{A}_S$ such that $\mathbf{A}_S\mathbf{x}_S^N = \mathbf{0}$. The vectors $\mathbf{x}_S - \mathbf{x}_S^N$, $\mathbf{x}_S + \mathbf{x}_S^N$ can then be constructed so that both satisfy (4.19a) of Lemma 2. Therefore, if $\mathbf{A}_S$ does not have full rank $\mathbf{x}_S$ (and consequently $\mathbf{x}_0$) will not be a unique solution. Finally, assume that the second condition of Theorem 8 does not hold, namely the vector $\mathbf{c}$ does not satisfy equations (4.39a) and (4.39b). This is contradictory to Lemma 2 and by strong duality of problem (4.5), it is concluded that conditions (4.19b) and (4.19c) will be violated. □

Similarly, for the case of $\ell_{\infty,1}$ norm minimization the following can be stated.

Theorem 9. $\mathbf{x}_0$ *is the unique optimum point of the primal linear program* (4.28) *if and only if a)* $\mathbf{A}_S$ *has full rank and b) there exists some* $\mathbf{c} \in \mathbb{R}^M$ *such that*

$$\mathbf{A}_j^T\mathbf{c} = \mathbf{v}_j, \quad j \in S \tag{4.41a}$$

$$\|\mathbf{A}_j^T\mathbf{c}\|_1 < 1, \quad j \notin S \tag{4.41b}$$

where $\mathbf{v}_j$ *is given as* $\mathbf{v}_j \in \mathbb{R}^d : \|\mathbf{v}_j\|_1 = 1 :$ $\begin{cases} v_l x_{0l} \geq 0 & \text{if} \quad |x_{0l}| = \|\mathbf{x}_{0j}\|_\infty \\ v_l = 0 & \text{otherwise} \end{cases}$.

Proof. Following the same principle as in Theorem 8, it is simple to prove Theorem 9. □

Suppose now that for $\mathbf{x}_0$ as defined in Theorems 8 and 9, a vector $\mathbf{v}$ exists containing the subdifferentials for $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization respectively. Then, an outcome of Theorems 8 and 9 is the following corollary.

Corollary 3. *Let $\mathbf{x}_0$ be a solution to the problem $\mathbf{Ax} = \mathbf{y}$ and $p = 2$ and $p = 1$ for the $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization problems respectively. If $\mathbf{A}_S$ has full rank and*

$$\left\|\mathbf{A}_j^T \mathbf{c}\right\|_p < 1, \quad \forall j \notin S \tag{4.42}$$

is satisfied with the dual vector $\mathbf{c} = (\mathbf{A}_S^\dagger)^T \mathbf{v}$, then $\mathbf{x}_0$ is the unique optimum.

Corollary 3 is for the group sparsity case the equivalent of Fuchs Corollary for the conventional sparsity model. Therefore, in a similar way as in the standard sparsity approach, condition (4.42) can be used in order to find recovery conditions for the GPFP algorithm. Note that although the conditions specified by Theorems 8 and 9 and Corollary 3 are general conditions for primal-dual uniqueness, it is not guaranteed that they will be satisfied by the proposed suite of algorithms, as explained in the previous section.

In what follows, the $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization problems are treated separately and stronger efficient recovery conditions than (4.42) for GPFP algorithms are provided. Interestingly, the results for GPFP with the $\ell_{2,1}$ objective function seem to be the same with the ones reported in [101] for the GOMP algorithm.

4.3.2 Efficient recovery conditions for $\ell_{2,1}$ minimization with GPFP

Define first the quantity $\rho_c(\mathbf{A})$ as in [101]:

$$\rho_c(\mathbf{A}) = \max_j \sum_i \rho(\mathbf{A}_{ij}) \tag{4.43}$$

where $\rho(\mathbf{A}) = \|\mathbf{A}\|_2 = \lambda_{\max}^{1/2}(\mathbf{A}^T\mathbf{A})$ is the spectral norm of matrix $\mathbf{A}$, $\lambda_{\max}$ is the maximum eigenvalue and $\mathbf{A}_{ij}$ denotes the (i,j)-th $d \times d$ block of $\mathbf{A}$. Define also the set $\bar{S}$ to be the complement of the true support set S. Then the following theorem can be stated.

Theorem 10. *Let $\mathbf{x}_0$ be a group sparse solution to the problem of* (4.5) *and suppose that $\mathbf{A}_S$ has full rank. If*

$$\rho_c(\mathbf{A}_S^\dagger \mathbf{A}_{\bar{S}}) < 1 \tag{4.44}$$

then $\mathbf{x}_0$ *is the unique solution to* (4.5).

Proof. In order to prove this, consider the inequality (4.42) from Corollary 3. The dual vector $\mathbf{c}$ is given by $\mathbf{c} = (\mathbf{A}_S^\dagger)^T\mathbf{v}$, where the vector $\mathbf{v}$ is the concatenation of the subdifferentials of the groups belonging to the true support set S. For any $j \notin S$ (or equivalently $j \in \bar{S}$), it holds:

$$\begin{aligned}
\|\mathbf{A}_j^T\mathbf{c}\|_2 &= \|\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T\mathbf{v}\|_2 \\
&\leq \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_2 \|\mathbf{v}_i\|_2 \\
&= \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_2 \|\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{2,1}\|_2 \\
&= \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_2 \|\mathbf{x}_i/\|\mathbf{x}_i\|_2\|_2 \\
&= \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_2 1 \\
&= \sum_i \|[\mathbf{A}_S^\dagger\mathbf{A}_j]_i\|_2
\end{aligned} \tag{4.45}$$

where the triangular inequality and equation (4.16) have been used. $[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i$ is the submatrix corresponding to the i-th group. It follows that:

$$\|\mathbf{A}_{\bar{S}}^T\mathbf{c}\|_{2,\infty} \leq \max_j \sum_i \|(\mathbf{A}_S^\dagger\mathbf{A}_j)_i\|_2 = \max_j \sum_i \rho([\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}]_{ij}) = \rho_c(\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}) \tag{4.46}$$

where $[\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}]_{ij}$ denotes the element corresponding to the i-th row and j-th column of the matrix. Combining now equations (4.46) and (4.42), it is straightforward to see that if $\rho_c(\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}) < 1$ is satisfied then (4.42) will always be satisfied. □

Although Theorem 10 provides a stronger condition than Corollary 3, it is intuitive in quantifying the robustness of GPFP algorithm with the $\ell_{2,1}$ constraint. First of all, note that (4.44) does not depend on the optimum dual vector $\mathbf{c}^*$. Furthermore, as can be seen from (4.45) the quantity $\sum_i \|(\mathbf{A}_S^\dagger\mathbf{A}_j)_i\|_2$ is an upper bound not only for the optimum vector $\mathbf{v}$, but for any vector $\mathbf{v}$ that spans the row space of $\mathbf{A}_S$ as long as it satisfies that $\|\mathbf{v}_i\|_2 = 1$ for any $i = 1, \ldots, K$. Subsequently, the condition (4.44) of Theorem 10 provides the recovery condition for GPFP using the $\ell_{2,1}$ constraint. Therefore, if (4.44) holds the proposed method will find the optimum solution vector $\mathbf{x}_0$.

The result of Theorem 10 was first introduced by Eldar et al. [101], where it was shown to be the efficient recovery condition for GOMP and mixed $\ell_{2,1}$ minimization. Furthermore, in the specific work, it was also shown that if (4.44) holds for an overcomplete dictionary $\mathbf{A}$ then the group sparsity level is bounded by:

$$Kd < \frac{1}{2}\Big(\mu_g^{-1} + d - (d-1)\frac{\mu_s}{\mu_g}\Big) \tag{4.47}$$

where d is the size of the group and μ_g and μ_s is the group coherence and sub-coherence of $\mathbf{A}$ respectively, given by:

$$\mu_g = \max_{1\leq i,i\neq j\leq \frac{N}{d}} \frac{1}{d}\rho(\mathbf{A}_i^T\mathbf{A}_j) \tag{4.48}$$

and

$$\mu_s = \max_{1\leq i\leq \frac{N}{d}} \max_{1\leq l,l\neq r\leq d} |\mathbf{a}_l^T\mathbf{a}_r|, \quad \mathbf{a}_l, \mathbf{a}_r \in \mathbf{A}_i \tag{4.49}$$

where the dictionary $\mathbf{A}$ is assumed to have normalized columns. It follows that as long as Theorem 10 is satisfied then the sparsity threshold given by (4.47) will also hold for GPFP. This result can be viewed as an extension of (2.42) to the group sparsity case. Indeed, if the size of group is set to be $d = 1$, the inequality $K < (\mu^{-1}+1)/2$ is obtained.

4.3.3 Efficient recovery conditions for $\ell_{\infty,1}$ minimization with GPFP

For the case of $\ell_{\infty,1}$ minimization, define the quantity $q_c(\mathbf{A})$, which is given by:

$$q_c(\mathbf{A}) = \max_j \sum_i \|\mathbf{A}_{ij}\|_\infty \tag{4.50}$$

where $\|\mathbf{A}\|_\infty = \max_{1\leq i\leq M}\sum_{j=1}^N A_{ij}$ is the infinity matrix norm, $\mathbf{A}_{ij}$ denotes the (i,j)-th $d\times d$ block of $\mathbf{A}$ and A_{ij} the (i,j)-th element of $\mathbf{A}$. Then the following theorem can be stated.

Theorem 11. *Let $\mathbf{x}_0$ be a group sparse solution to the problem of* (4.28) *and suppose that $\mathbf{A}_S$ has full rank. If*

$$q_c(\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}) < 1 \tag{4.51}$$

then $\mathbf{x}_0$ is the unique solution to (4.28).

Proof. To prove this, consider the inequality (4.42) from Corollary 3. The dual vector $\mathbf{c}$ is given by $\mathbf{c} = (\mathbf{A}_S^\dagger)^T\mathbf{v}$, where vector $\mathbf{v}$ is the concatenation of the subdifferentials $\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{\infty,1}$ of the groups belonging to the true support set S. For any $j \notin S$ (or equivalently $j \in \bar{S}$), it holds:

$$\begin{aligned}
\|\mathbf{A}_j^T\mathbf{c}\|_1 &= \|\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T\mathbf{v}\|_1 \\
&\leq \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_1 \|\mathbf{v}_i\|_1 \\
&= \sum_i \|(\mathbf{A}_j^T[\mathbf{A}_S^\dagger)^T]_i\|_1 \|\partial_{\mathbf{x}_i}\|\mathbf{x}_i\|_{\infty,1}\|_1 \\
&= \sum_i \|[\mathbf{A}_j^T(\mathbf{A}_S^\dagger)^T]_i\|_1 1 \\
&= \sum_i \|[\mathbf{A}_S^\dagger\mathbf{A}_j]_i\|_\infty
\end{aligned} \tag{4.52}$$

where the triangular inequality and equation (4.32) have been used. It follows that:

$$\|\mathbf{A}_{\bar{S}}^T\mathbf{c}^*\|_{1,\infty} \leq \max_j \sum_i \|(\mathbf{A}_S^\dagger\mathbf{A}_j)_i\|_\infty = \max_j \sum_i \|[\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}]_{ij}\|_\infty = q_c(\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}). \tag{4.53}$$

Combining now (4.53) and (4.42), it can be shown that as long as $q_c(\mathbf{A}_S^\dagger\mathbf{A}_{\bar{S}}) < 1$ is satisfied (4.42) will be also satisfied. □

Similar to case of $\ell_{2,1}$ minimization, Theorem 11 provides the theoretical guarantees for the GPFP algorithm with the $\ell_{\infty,1}$ constraint. The inequality (4.51) will hold for any vector $\mathbf{v}$ that spans the row space of $\mathbf{A}_S$ and satisfies $\|\mathbf{v}_i\|_1 = 1$. Therefore, considering that GPFP does not guarantee convergence to the optimum vector $\mathbf{c}^*$, if condition (4.51) is satisfied then the algorithm will converge to the optimum solution $\mathbf{x}_0$.

Comparing Theorems 10 and 11, which give the efficient recovery conditions for $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization using the GPFP algorithm, it can be observed that the spectral norm in (4.43) is replaced by the $\|.\|_\infty$ matrix norm in (4.50). Therefore, considering the fact that for the spectral and infinity norms of a matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ the inequality $\frac{1}{\sqrt{N}}\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{M}\|\mathbf{A}\|_\infty$ holds, it is expected that the recovery performance of the two approaches will differ from a theoretical point of view. In Section 4.3.5, the algorithms' performance is examined using an average case analysis.

4.3.4 Interpretation of GPFP recovery conditions

Consider once again the geometry of the dual problem associated with the primal linear program for both conventional and group sparsity models. For simplicity and without loss of generality, consider the case of sparsity $K = 2$. When no assumptions on the structure of the coefficients vector are made and sparsity is enforced by the ℓ_1 norm, the feasible points for the dual vector $\mathbf{c}$ are the vertices of the polar polytope corresponding to the dictionary atoms that belong to the true support set (i.e. $\mathbf{a}_i \in \mathbf{A}_S$). As can be seen in Fig. 4.5, in the special case of nonnegative coefficients the vertex $\mathbf{c}_1$ is the only feasible solution for the dual vector $\mathbf{c}$. However, when the group structure is considered the corresponding optimum vector $\mathbf{c}$ can lie anywhere on the first quadrant of the unit circle for the problem of mixed $\ell_{2,1}$ norm minimization. On the other hand, if the $\ell_{\infty,1}$ norm is set to be the objective function, it turns out that there are only three feasible solutions for $\mathbf{c}$ in the nonnegative setting, as illustrated in Fig. 4.5. This can be easily verified by taking into account condition (4.41a) of Theorem 9. Note that the ℓ_p balls shown in Fig. 4.5, coincide with the duals to the norms of the objective function for each case.

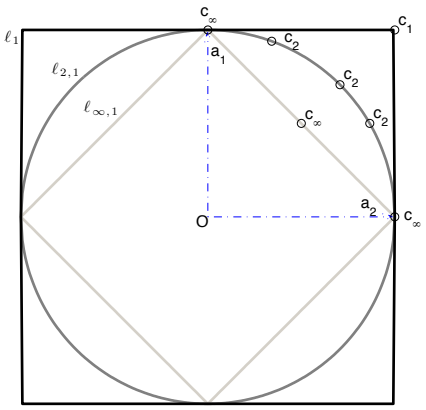


Figure 4.5: Dual balls and feasible points for the dual vector $\mathbf{c}$ in the case of ℓ_1, $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization. As can be seen, in the nonnegative setting (corresponding to the first quadrant of the cartesian coordinate system), there is only one feasible point for ℓ_1 minimization, three feasible points for $\ell_{\infty,1}$ minimization, while for $\ell_{2,1}$ minimization $\mathbf{c}$ can lie at any point on the unit ℓ_2 ball.

Therefore, one way to interpret the results of Theorems 10 and 11 for the case of $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization from a geometrical point of view is to deduce that as long as the corresponding conditions (4.44) and (4.51) hold, then GPFP for any feasible vector $\mathbf{c}_2$ and $\mathbf{c}_\infty$ will identify the true support set.

4.3.5 Phase transitions of GPFP algorithms

In this section, the performance of the proposed method is quantified by empirically measuring the success of GPFP algorithms in group sparse recovery problems. To do this, synthetic group sparsity problems were generated for variable over-sampling (M/N) and under-sampling (Kd/M) ratios, which were allowed to be in the range of [0, 1]. The dictionary columns were drawn from the random uniform ensemble and their dimension was set at $N = 256$. They were also clustered in groups of equal size $d = 4$ atoms. For each problem instance, namely each set of parameters M, N and K the experiment was repeated 100 times.

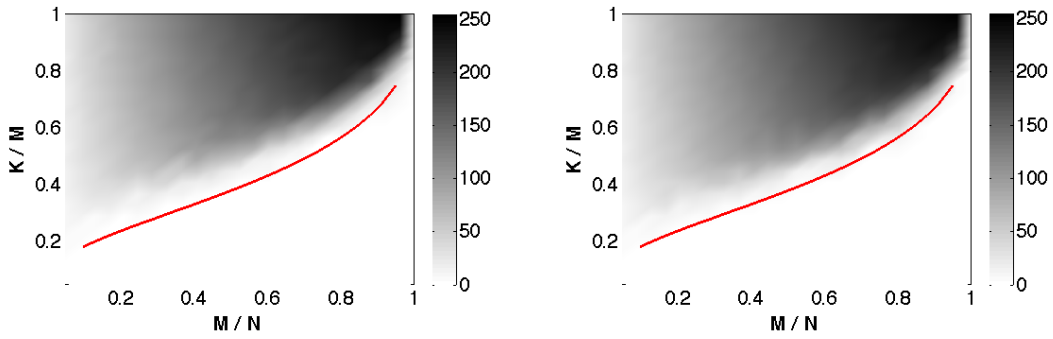


Figure 4.6: Phase diagrams of (a) GPFP with $\ell_{2,1}$ constraint and (b) GPFP algorithm with $\ell_{\infty,1}$ constraint.

Fig. 4.6(a)-(b) illustrate the resulting phase diagrams for GPFP with $\ell_{2,1}$ and $\ell_{\infty,1}$ constraints respectively. The shaded area is the number of coefficients that differ more than 10^{-4} from the optimal sparse solution. Both algorithms display a significantly improved performance over ℓ_1 minimization and subsequently the original conventional sparsity PFP algorithm. As can be seen, the phase transitions, in both cases, are shifted notably above the theoretical red curve, corresponding to ℓ_1 minimization. Between the two GPFP algorithms, it can be observed that GPFP based on the $\ell_{2,1}$ constraint shows a slightly better performance, achieving superior phase transition in the examined case with the group size being equal to $d = 4$.

4.4 Experimental results for the group sparsity model

This section presents further experimental results based on synthetic data that demonstrate the group sparse recovery capability of the GPFP algorithm . Its performance is

also compared against SparseLab's implementation of PFP and other existing methods that incorporate the group sparsity structure such as Block Orthogonal Matching Pursuit (BOMP) and Group Orthogonal Matching Pursuit (GOMP), as implemented in the Matlab toolbox GroupSparseBox. The main difference between BOMP and GOMP appears at the group of atoms selection criterion step, for which and according to this implementation BOMP selects a new group based on the highest correlation of each group while GOMP's group selection is based on largest ℓ_2 norm of residual correlations of each group. Regarding the proposed method, both versions of GPFP with the $\ell_{2,1}$ and $\ell_{\infty,1}$ objective functions are considered.

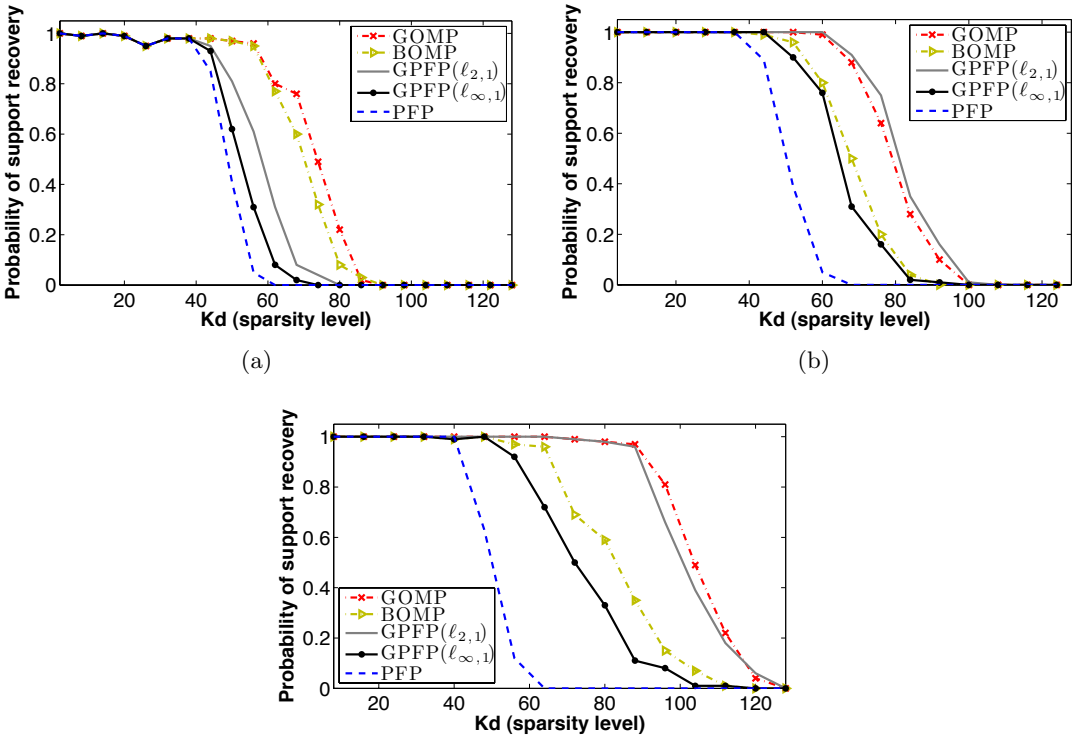


Figure 4.7: Support recovery rates (over 100 trials) of GOMP, BOMP, GPFP (with $\ell_{2,1}$ and $\ell_{\infty,1}$ constraints) and PFP vs sparsity level Kd for a dictionary $\mathbf{A} \in \mathbb{R}^{M\times N}$ with $M = 128$, $N = 256$ and group size (a) $d = 2$, (b) $d = 4$ and (c) $d = 8$.

For the first experiment, dictionaries of size 128×256 with columns drawn from an i.i.d. zero-mean Gaussian distribution were randomly generated. The block K-sparse vector $\mathbf{x}$ with block size d was generated by selecting uniformly at random the support of the nonzero groups of atoms and filling the corresponding values by sampling from the normal distribution. The block sparsity level K was variable ranging from 1 to M/d. Fig. 4.7(a)-(c) show the average support recovery rates over 100 iterations for all

tested algorithms against the sparsity level Kd for $d = 2$, $d = 4$ and $d = 8$, respectively. The probability of support recovery stands for the proportion of the 100 sparse vectors that are exactly recovered. For exact recovery, it is required that the support of the recovered vector $\hat{\mathbf{x}}$ perfectly matches that of the true solution vector $\mathbf{x}$. Hence, if a single false detection occurs due to a wrong atom selection, it is considered that the tested algorithm has failed. As can be seen, in all plots for sparsity levels less than $Kd = 40$ all algorithms achieve the same recovery performance, whereas for sparsity levels $Kd > 40$ it can be noticed that the group sparsity algorithms perform better in all cases and this performance gain increases with the block size. Among the group sparsity versions of OMP and PFP, BOMP and GOMP outperform both GPFP algorithms in the case when the group size is $d = 2$, with GOMP achieving the best overall recovery rates. However, this performance gain vanishes as the block size becomes larger. In fact, for group sizes $d = 4$ the GPFP algorithm based on $\ell_{2,1}$ minimization has a slight edge in terms of performance over GOMP while outperforming all other algorithms tested. The performance of GPFP based on the $\ell_{\infty,1}$ minimization also improves, as the specific algorithm achieves recovery rates very close to the ones obtained for the BOMP algorithm. Although it still displays the worst performance among the group sparsity methods, its performance gain over the conventional PFP algorithm cannot be considered negligible.

In Fig. 4.8(a)-(c), the average elapsed times over 100 iterations are plotted in log-scale for all algorithms tested against the sparsity level Kd for group sizes $d = 2$, $d = 4$ and $d = 8$ respectively. As can be seen, in all cases BOMP achieves the fastest convergence. From the remaining algorithms, when $d = 2$ GOMP is the fastest for sparsity levels $Kd > 20$, while for $Kd < 20$ PFP shows the fastest convergence. As expected, due to the additional computational cost in the group selection criterion step, the proposed GPFP methods are the slowest for the specific setting. For instance, when the sparsity level is $Kd = 56$ BOMP converges after 1.8 sec, GOMP after 8.4 sec and PFP after 26.1 sec. However, the average elapsed time values corresponding to GPFP ($\ell_{2,1}$) and GPFP ($\ell_{\infty,1}$) are 28.3 sec and 61.4 sec, respectively. Nevertheless, as the group size increases their convergence speed decreases. As illustrated in Fig. 4.8(a)-(c), for group sizes $d = 8$ and sparsity levels $Kd > 50$ the average elapsed time for GPFP ($\ell_{2,1}$) is less than the time needed for the convergence of GOMP. GPFP ($\ell_{\infty,1}$) is somehow slightly slower than

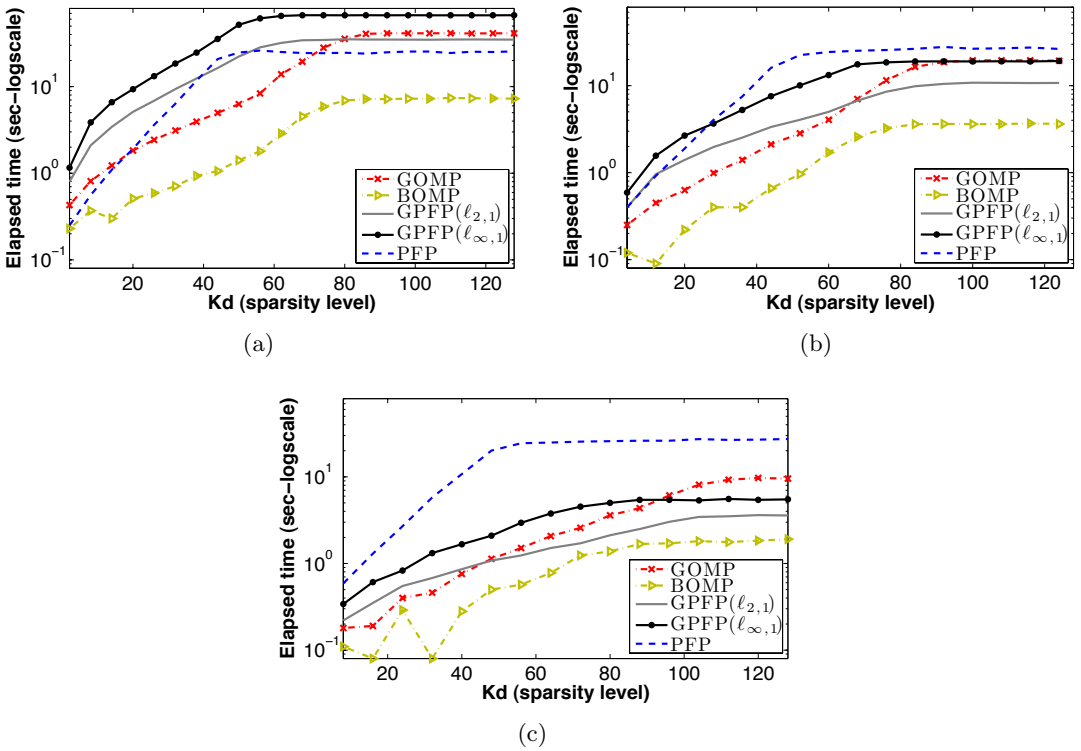


Figure 4.8: Average elapsed time (in log-scale) over 100 trials for GOMP, BOMP, GPFP (with $\ell_{2,1}$ and $\ell_{\infty,1}$ constraints) and PFP vs sparsity level Kd for a dictionary $\mathbf{A} \in \mathbb{R}^{M\times N}$ with $M = 128$, $N = 256$ and group size (a) $d = 2$, (b) $d = 4$ and (c) $d = 8$

the rest of the group sparsity algorithms, however when compared to the original PFP algorithm the gain in computational time is evident. For the value of sparsity $Kd = 56$, BOMP is the fastest algorithm requiring an average time of 0.6 sec to converge, followed by GPFP ($\ell_{2,1}$) with an average elapsed time of 1.2 sec. The convergence times for GOMP and GPFP ($\ell_{\infty,1}$) in that case are 1.5 sec and 2.5 sec, respectively. PFP achieves the slowest convergence with an average time of 24.5 sec. Therefore, apart from the increase in performance gain the group sparse algorithms can also reduce the computational time as opposed to the standard sparsity approach when the group size becomes larger.

Fig. 4.9 presents simulation results for the same experimental setting as above, but using a four times overcomplete dictionary $\mathbf{A} \in \mathbb{R}^{64\times 256}$. Fig. 4.9(a)-(b) illustrate the recovery rates and the average elapsed times respectively of all algorithms tested for groups of size $d = 2$. Fig. 4.9(c)-(d) show the analogous results for the case of $d = 4$. It can be observed that all results for the four times overcomplete dictionary are in almost full agreement with the ones obtained in the first experiment using a twice overcomplete dictionary.

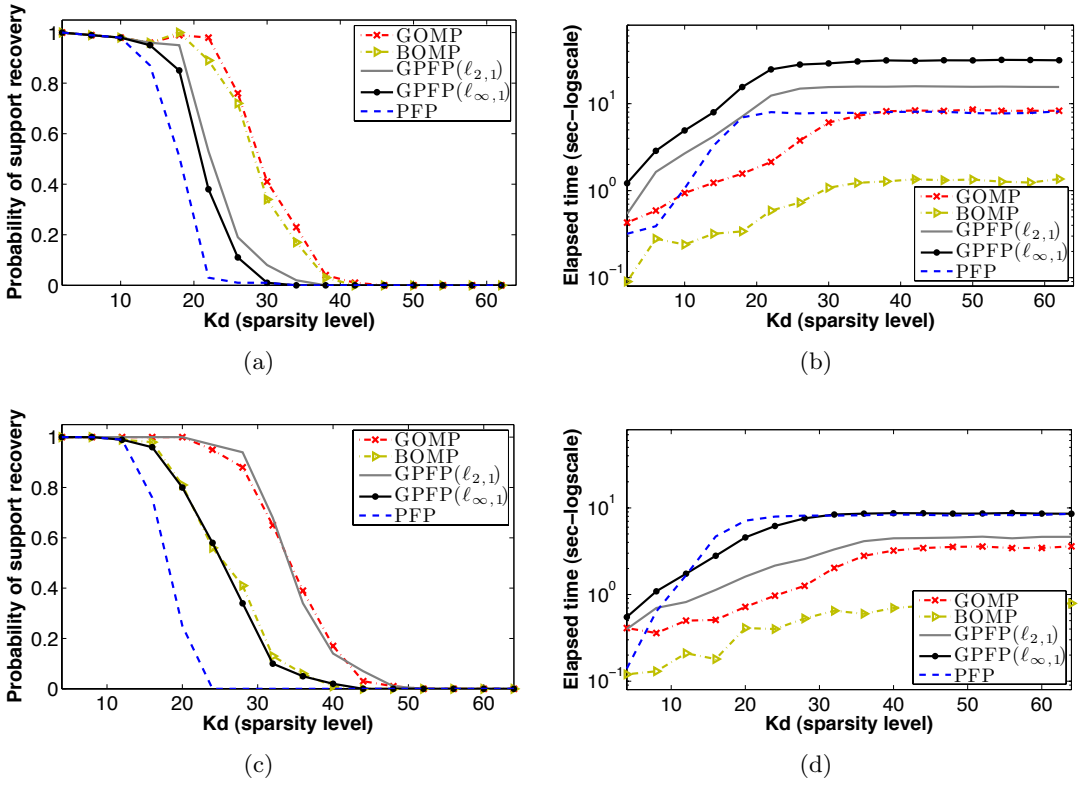


Figure 4.9: Support recovery rates (over 100 trials) and average elapsed times of GOMP, BOMP, GPFP (with $\ell_{2,1}$ and $\ell_{\infty,1}$ constraints) and PFP vs sparsity level Kd for a dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M = 64$, $N = 256$ and group sizes $d = 2$ ((a)-(c)) and $d = 4$ ((b)-(d)).

In the previous tests, the sparse vector $\mathbf{x}$ was synthesized by sampling from the normal distribution, which generally favours greedy algorithms such as MP and OMP in the conventional sparsity case [113]. On the other hand, the least favourable distribution appears to be the equiprobable distribution. Based on that the last experiment attempts to investigate the performance of all tested algorithms in the scenario when the entries of the sparse vector $\mathbf{x}$ follow the equiprobable distribution in $\{-1, 1\}$. This was done by keeping the same setting as in the first experiment and using a twice overcomplete dictionary of size 128×256. As can be seen in Fig. 4.10(a) for group sizes $d = 2$, the proposed methods achieve the best recovery rates. However, this performance gain diminishes as the group size increases and especially GPFP ($\ell_{\infty,1}$) exhibits similar shortcomings as in the previous tests (Fig. 4.10(b)-(c)). Nevertheless, GPFP ($\ell_{2,1}$) achieves the best overall performance for almost all problem suites examined.

Regarding the GPFP algorithms, GPFP with the $\ell_{2,1}$ exhibits the best performance in all cases and for all group sizes, which is in agreement with the phase transitions shown

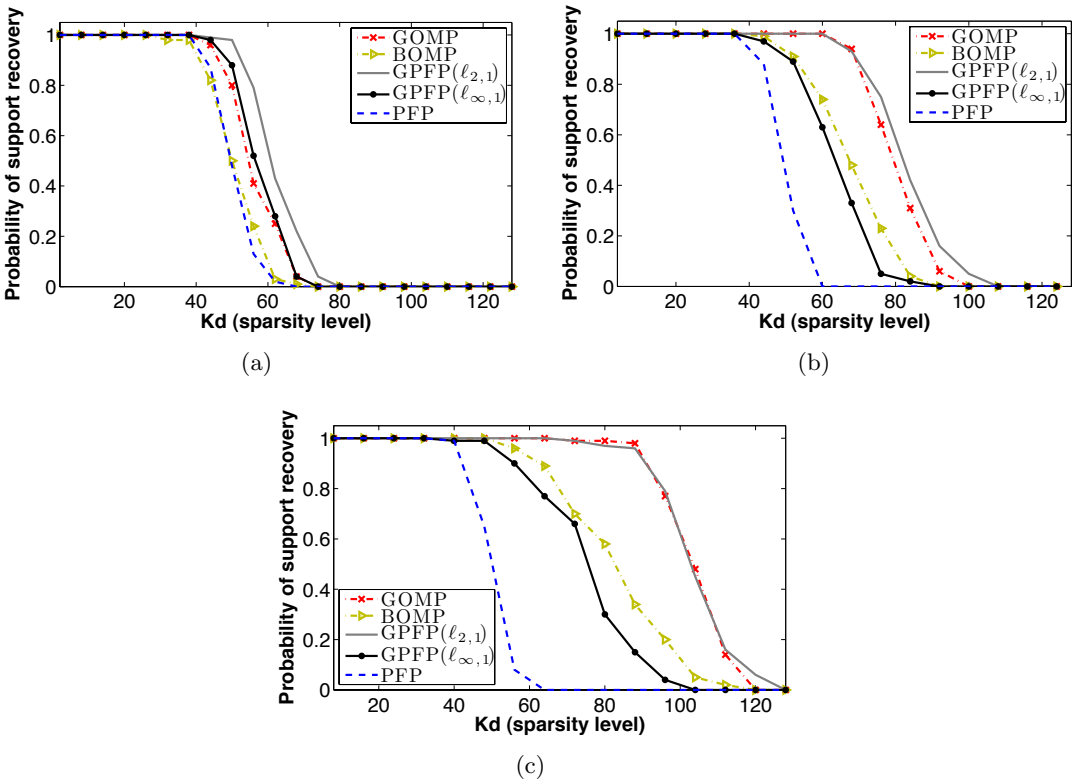


Figure 4.10: Support recovery rates (over 100 trials) of GOMP, BOMP, GPFP (with $\ell_{2,1}$ and $\ell_{\infty,1}$ constraints) and PFP vs sparsity level Kd for a dictionary $\mathbf{A} \in \mathbb{R}^{M\times N}$ with $M = 128$, $N = 256$ and group size (a) $d = 2$, (b) $d = 4$ and (c) $d = 8$.

in Fig. 4.6.

4.5 Joint sparse recovery with PFP

All previous sections of the current chapter presented efficient recovery algorithms for the problems of sparse representations and compressed sensing. In this discussion, the single measurement vector (SMV) sparsity model was considered, implying that the observations are given by the single channel vector $\mathbf{y} \in \mathbb{R}^M$. In compressed sensing this measurement vector is obtained by projecting the discrete K-sparse high dimensional vector $\mathbf{x} \in \mathbb{R}^N$ to a lower M dimensional space using random projections. It has been already discussed that the main assumption for efficient recovery is that $\mathbf{x}$ is sufficiently sparse or group sparse. However, in many practical signal processing applications emerging in fields such as array processing [3, 114] or magnetoencephalography [96, 115, 116], multiple measurement vectors, corresponding to multiple unknown vectors that share a common sparse support, might be encountered. The specific setting can be viewed as

another type of structure where rather than attempting to recover a single sparse vector $\mathbf{x}$ at a time, the aim is to simultaneously recover all jointly sparse vectors that comprise the multiple measurements. Therefore, a natural extension of the SMV model is the so-called multiple measurement vector (MMV) model [117], which addresses the underlying joint sparse recovery problem. It has been shown in [60] that this problem formulation attempts to simultaneously recover the unknown signal support set exploiting this joint sparsity in the sparsity pattern and it is a special case of the group structure. More specifically, it can be proven that any MMV sparse recovery problem can be reformulated as a group sparse vector recovery problem by appropriately interleaving the measurements matrix and constructing the corresponding single vector of measurements [98]. However, the opposite does not always hold.

Suppose now that $\mathbf{Y} \in \mathbb{R}^{M \times L}$ is a collection of L measurement vectors $\mathbf{y}_{i=1}^{L}$ that are obtained through the common compressed sensing matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ and $\mathbf{X} \in \mathbb{R}^{N \times L}$ is the jointly K-sparse matrix. In other words, it is assumed that there are at most K rows in $\mathbf{X}$ containing nonzero values. Therefore, the assumption of the MMV model is that $\|\mathbf{X}\|_0 = |\text{supp}(\mathbf{X})| \leq K$, where $|\text{supp}(\mathbf{X})|$ is the size of the support set $\text{supp}(\mathbf{X})$ which is defined as in [118] to be:

$$\text{supp}(\mathbf{X}) = \bigcup_{j=1}^{L} \text{supp}(\mathbf{x}_j) \tag{4.54}$$

with the support of each individual column be given by $\text{supp}(\mathbf{x}_j) = \{i, x_{ij} \neq 0\}$, where x_{ij} corresponds to the i-th element of the j-th column of matrix $\mathbf{X}$. Therefore, in the noiseless case the joint sparse recovery problem is to identify the row support of the unknown jointly sparse matrix $\mathbf{X}$ given the MMV matrix $\mathbf{Y}$ defined as:

$$\mathbf{Y} = \mathbf{A}\mathbf{X}. \tag{4.55}$$

After determining the unknown support set, the nonzero entries of the sparse matrix $\mathbf{X}$ can be recovered by a Moore-Penrose inversion. In a similar way to the SMV problem, the MMV sparse recovery problem can be cast as a constrained optimization problem:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_0 \quad \text{such that} \quad \mathbf{Y} = \mathbf{A}\mathbf{X}. \tag{4.56}$$

However, the MMV ℓ_0 minimization problem of (4.56), as happens to be in the SMV case, is combinatorial (e.g. NP-hard) [67]. Nevertheless, several suboptimal approaches with polynomial complexity have been proposed that under certain conditions can recover the unknown jointly sparse matrix $\mathbf{X}$. The authors in [119] propose the mixed $\ell_{2,1}$ minimization problem defined as:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{2,1} \quad \text{such that} \quad \mathbf{Y} = \mathbf{A}\mathbf{X} \tag{4.57}$$

where $\|\mathbf{X}\|_{2,1} = \sum_{i=1}^{N} \|\mathbf{X}_i\|_2$ and $\mathbf{X}_i$ denotes the i-th row of $\mathbf{X}$. The minimization problem of (4.57) is convex and can be tackled using interior-point methods (or SOCP methods). This can be viewed as an extension of the BP algorithm to the multiple vectors scenario, since when $L = 1$ the problem is reduced to the standard ℓ_1 minimization problem. As in the case of union of subspaces and group structured sparsity, greedy algorithms such as OMP have also been extended to the MMV problem providing faster alternatives to (4.57). The so-called Simultaneous Orthogonal Matching Pursuit (SOMP), is an iterative greedy algorithm, which operates adding one dictionary atom at a time that has been identified according to the criterion:

$$i^k = \arg \max_{i \notin \mathcal{I}^{k-1}} \|\mathbf{a}_i^T \mathbf{R}^{k-1}\|_2 \tag{4.58}$$

where $\mathbf{R}^{k-1}$ is the current residual, initialized as $\mathbf{R}^0 = \mathbf{Y}$. Similar to OMP, SOMP proceeds with updating the residual matrix $\mathbf{R}$ and the solution matrix $\mathbf{X}$. The whole process is repeated until there is no residual left.

Theoretical guarantees of greedy algorithms and convex optimization for the MMV problem have been studied extensively in recent literature [95, 120, 98, 119]. Although these theoretical results suggest equivalence between MMV and SMV models showing no performance gain, in practice and according to empirical results both $\ell_{2,1}$ minimization and SOMP can vastly improve the recovery ability of the conventional sparsity approaches.

4.5.1 Dual linear program of the MMV problem & MMV-PFP algorithm

The Lagrangian to the problem of equation (4.57) is given by:

$$\mathcal{L}(\mathbf{x}, \mathbf{c}) = \|\mathbf{X}\|_{2,1} - \mathbf{C}^T(\mathbf{A}\mathbf{X} - \mathbf{Y}). \tag{4.59}$$

After minimizing the Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{c})$, it is straightforward to show that the dual to the MMV problem is given by:

$$\max_{\mathbf{C}} \operatorname{tr}(\mathbf{Y}^T\mathbf{C}) \quad \text{such that} \quad \|\mathbf{A}^T\mathbf{C}\|_{2,\infty} \leq 1. \tag{4.60}$$

The KKT conditions require that $\mathbf{Y} = \mathbf{A}\mathbf{X}$ and $\partial_{\mathbf{X}}\mathcal{L}(\mathbf{X}, \mathbf{C}) \in \mathbf{0}$. A similar analysis as presented in Section 4.2.1 for the group sparsity problem, yields the following lemma, which provides the necessary and sufficient conditions for the optimum primal-dual pair $\mathbf{X}^*$ and $\mathbf{C}^*$.

Lemma 4. *Suppose that the primal problem of equation* (4.57) *is feasible. Then the pair* $\mathbf{X}^*$, $\mathbf{C}^*$ *is an optimum solution pair for both primal and dual linear programs if and only if the following conditions hold:*

$$\mathbf{A}\mathbf{X}^* = \mathbf{Y} \tag{4.61a}$$

$$\|\mathbf{A}^T\mathbf{C}^*\|_{2,\infty} \leq 1 \tag{4.61b}$$

$$\mathbf{Y}^T\mathbf{C}^* = \|\mathbf{X}\|_{2,1}. \tag{4.61c}$$

Proof. This follows immediately from KKT conditions. □

Lemma 4 suggests that once again the underlying geometry can be exploited to derive an algorithm for the MMV problem that promotes joint sparsity across the rows of $\mathbf{X}$. Adopting a similar approach as described in Section 4.2.2 for the group structure, at the k-th step the algorithm requires that $\mathbf{a}_i^T(\mathbf{C}^{k-1} + \lambda\mathbf{R}^{k-1}) = \mathbf{v}_i^k$, where it is chosen that $\mathbf{v}_i^k = \mathbf{a}_i^T\mathbf{R}^{k-1}/\|\mathbf{a}_i^T\mathbf{R}^{k-1}\|_2$ such that $\|\mathbf{v}_i^k\|_2 = 1$. It turns out that the atom selection criterion for the MMV-Polytope Faces Pursuit (MMV-PFP) algorithm is given by:

$$\alpha^k = \arg\max_{i \notin \mathcal{I}^{k-1}} \frac{\|\mathbf{a}_i^T\mathbf{R}^{k-1}\|_2^2}{\|\mathbf{a}_i^T\mathbf{R}^{k-1}\|_2 - \mathbf{a}_i^T\mathbf{C}^{k-1}(\mathbf{R}^{k-1})^T\mathbf{a}_i}. \tag{4.62}$$

where $\mathbf{R}^k$ denotes the residual at the k-th step of the algorithm. It can be easily verified that when the number of channels is $L = 1$ equation (4.62) reduces to the maximum scaled correlation of equation (3.7). The resulting algorithm of MMV-PFP is given in Algorithm 8.

Algorithm 8 MMV-Polytope Faces Pursuit

Input: $\mathbf{A}$, $\mathbf{Y}$
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Initialize:
$k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{C}^k \leftarrow \mathbf{0},\ \mathbf{X}^k \leftarrow \emptyset,\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{0},\ \mathbf{R}^k \leftarrow \mathbf{Y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \mathbf{a}_i^T \mathbf{R}^{k-1} > \theta_{\min}$ **do** {Find next face}
  $k \leftarrow k + 1$
  Find face:
  $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \{\|\mathbf{a}_i^T \mathbf{R}^{k-1}\|_2^2 / (\|\mathbf{a}_i^T \mathbf{R}^{k-1}\|_2 - \mathbf{a}_i^T \mathbf{C}^{k-1} (\mathbf{R}^{k-1})^T \mathbf{a}_i)\}$
  Add constraints:
  $\mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{a}_{i^k}],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\mathbf{X}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{Y},\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{A}^k \mathbf{X}^k,\ \mathbf{R}^k \leftarrow \mathbf{Y} - \hat{\mathbf{Y}}^k$
  $\mathbf{v}^k = [\mathbf{v}^{k-1}, \mathbf{a}_{i^k}^T \mathbf{R}^{k-1} / \|\mathbf{a}_{i^k}^T \mathbf{R}^{k-1}\|_2],\ \mathbf{C}^k \leftarrow [(\mathbf{A}^k)^\dagger]^T \mathbf{v}^k$
end while
Output: $\mathbf{X}^* = \mathbf{X}^k$

4.5.2 Uniqueness & recovery conditions

The following theorem, based on Lemma 4 provides the necessary and sufficient conditions for $\mathbf{X}_0$ to be the unique optimum solution of problem (4.57).

Theorem 12. *$\mathbf{X}_0$ is the unique optimum point of the primal MMV linear program if and only if a) $\mathbf{A}_S$ has full rank and b) there exists some $\mathbf{C} \in \mathbb{R}^{M \times L}$ such that*

$$\mathbf{a}_j^T \mathbf{C} = \mathbf{X}_{0j} / \|\mathbf{X}_{0j}\|_2, \quad \forall j \in S \tag{4.63a}$$

$$\|\mathbf{a}_j^T \mathbf{C}\|_2 < 1, \quad \forall j \notin S. \tag{4.63b}$$

Proof. The proof is similar to the one for Theorem 8. □

Define now the matrix $\mathbf{v}$ containing the subdifferentials of a potential solution $\mathbf{X}_0$ to the MMV problem. An immediate consequence of Theorem 8 is the following corollary:

Corollary 4. *Let $\mathbf{X}_0$ be a solution to the MMV problem $\mathbf{AX} = \mathbf{Y}$. If $\mathbf{A}_S$ has full rank and*

$$\|\mathbf{a}_j^T \mathbf{C}\|_2 < 1, \quad \forall j \notin S \tag{4.64}$$

is satisfied with the dual vector $\mathbf{C} = (\mathbf{A}_S^{\dagger})^T\mathbf{v}$*, then* $\mathbf{X}_0$ *is the unique optimum.*

Corollary 4 provides a uniqueness recovery condition (4.64) for the MMV mixed norm minimization problem. Nevertheless, following a worst case analysis in order to find a more handy condition than (4.64), one arrives at the result of the single channel problem, namely the ERC of (2.44). To see this, consider (4.64). For any $j \notin S$, it holds:

$$\begin{aligned}\left\|\mathbf{a}_j^T\mathbf{C}\right\|_2 &= \left\|\mathbf{a}_j^T(\mathbf{A}_S^{\dagger})^T\mathbf{v}\right\|_2 \\ &\leq \sum_i |[\mathbf{a}_j^T(\mathbf{A}_S^{\dagger})^T]_i| \|\mathbf{v}_i\|_2 \\ &= \sum_i |[\mathbf{a}_j^T(\mathbf{A}_S^{\dagger})^T]_i| 1 \\ &= \left\|\mathbf{A}_S^{\dagger}\mathbf{a}_j\right\|_1. \end{aligned} \tag{4.65}$$

Equation (4.65) is quite pessimistic as it shows no performance gain between the SMV and MMV versions of PFP algorithm. Interestingly, similar results exist for other algorithms such as SOMP and mixed $\ell_{2,1}$ minimization under this worst case analysis approach. Actually, the worst case analysis assumes that the multiple jointly sparse vectors will contain the same discrete signals and in that case, it is clear that the MMV model will not provide additional information and consequently no advantage in comparison to the SMV problem. This observation is in total agreement with Chen and Huo's elegant result, summarized in the following theorem.

Theorem 13. (Chen and Huo [120]) *A sufficient condition for the measurements* $\mathbf{Y} = \mathbf{AX}$ *to uniquely determine the jointly sparse matrix* $\mathbf{X}$ *is*

$$K = |\text{supp}(X)| < \frac{\text{spark}(\mathbf{A}) - 1 + \text{rank}(\mathbf{Y})}{2}. \tag{4.66}$$

It is noted that in the worst case scenario, where $\text{rank}(\mathbf{X}) = 1$, it is $\text{rank}(\mathbf{Y}) = 1$ since $\text{rank}(\mathbf{Y}) \leq \text{rank}(\mathbf{X})$ and therefore (4.66) reduces to the single channel sparsity threshold as given in equation (2.39) of Theorem 1 (Chapter 2). Nevertheless, the worst case scenario hypothesis might be a very restrictive and unrealistic assumption for practical applications. For instance, in the problem of source localization and DOA estimation, the sources to be localized in most cases happen to vary in the time or time-frequency domain. As will be seen in the following chapter, in this case $\mathbf{X}$ is a jointly sparse

matrix and its nonzero rows, corresponding to the location of the sources to be localized, contain the signals of interest. Consequently, it is quite unlikely that all channels of the sparse matrix $\mathbf{X}$ will contain the same exact information. Hence, an average case analysis could give more insight to the practical advantages of the MMV model. Indeed, as shown in [121, 119] an average case analysis not only explains the performance gain of the MMV model over the SMV but it also provides better recovery conditions under certain probabilistic assumptions.

Another important asset of Theorem 13 is that it clearly indicates what the worst case theoretical analysis fails to reveal. Inequality 4.66 suggests that as long as $\text{rank}(\mathbf{Y}) > 1$, there will be notable benefits from the exploitation of the joint sparsity property. One crucial benefit in that case will be the fact that the threshold on the sparsity level K will be increased. Thus, it is expected that the MMV algorithms will lead to enhanced recovery performance, allowing less sparse signals to be perfectly recovered. In the best case scenario, it is obvious that the rows of the sparse matrix $\mathbf{X}$ that correspond to its support set $\text{supp}(\mathbf{X})$ will be linear independent and hence the rank of $\mathbf{X}$ will be equal to the cardinality of the support, namely $\text{rank}(\mathbf{X}) = K$. From its definition, it hods that $\text{spark}(\mathbf{A}) \leq \text{rank}(\mathbf{A}) + 1 \leq M + 1$. Assuming that $\mathbf{A}$ is a full spark and well conditioned matrix the equality will hold. Since $\text{rank}(\mathbf{Y}) \leq \text{rank}(\mathbf{X})$, in the best case scenario the sparsity level is bounded by:

$$K < (M + K)/2 \Rightarrow K < M \Rightarrow K \leq M - 1. \tag{4.67}$$

The sparsity threshold of equation (4.67) admits a great improvement in the performance of the MMV model over the SMV. It suggests that for a given measurement matrix $\mathbf{Y} \in \mathbb{R}^{M \times L}$, the MMV model can recover sparse vectors with sparsity levels up to $M - 1$ as opposed to the value $M/2$ of the SMV model. On the other hand, for a given sparsity level K inequality (4.67) shows that from a compressed sensing point of view the length of measurement vectors M can be decreased, achieving further dimensionality reduction.

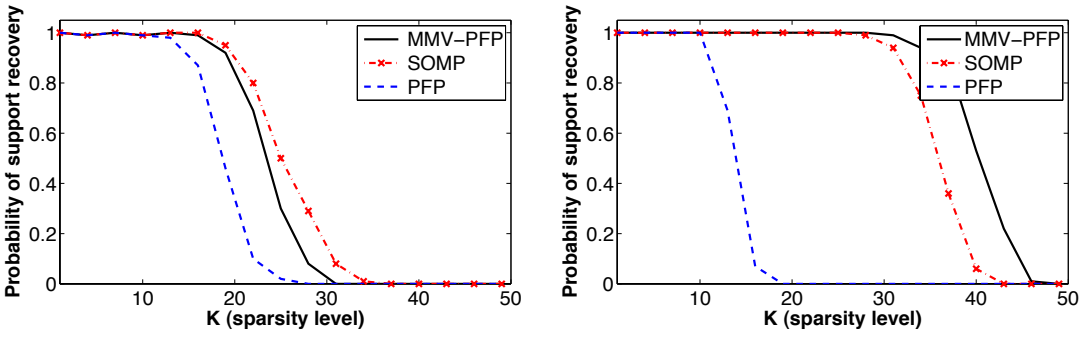
To conclude, although the theoretical guarantees for the proposed MMV-PFP algorithm are worse than the ones obtained for the standard PFP algorithm, it is expected that in practical application scenarios, i.e. in DOA estimation, where the use of multiple

channels is usually concerned with additional information, MMV-PFP will be capable of exploiting this additional information providing better performance recovery of the unknown support of $\mathbf{X}$.

4.6 Experimental results for the MMV model

In this section, the performance of MMV-PFP in the scenario of joint sparsity is evaluated and its recovery capabilities are compared against standard PFP and SOMP algorithms.

Fig. 4.11(a)-(b) show plots of the support recovery rate achieved by MMV-PFP, SOMP and PFP as a function of the sparsity level K when the number of the measurement vectors is $L = 2$ and $L = 50$ respectively. In the case of PFP, each single channel sparse recovery problem was solved separately. The problem size was chosen to be $M = 50$ and $N = 100$ and the corresponding dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ was generated by drawing its columns from the normal distribution. The jointly sparse matrix $\mathbf{X}$ was synthesized by selecting K rows at random each time and filling the corresponding nonzero entries with values drawn from the normal distribution. All results have been averaged over 100 times. As shown in all plots, although the theoretical guarantees for MMV-PFP are no better than PFP for the SMV model, MMV-PFP displays a significant gain in recovery performance, which increases with the number of measurements. When compared to SOMP, MMV-PFP shows better performance for $L = 50$, but slightly worse for $L = 2$.



(a) (b)

Figure 4.11: Support recovery rates (over 100 trials) of SOMP, MMV-PFP and PFP vs sparsity level K for a dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M = 50$, $N = 100$ and number of channels (a) $L = 2$ and (b) $L = 50$.

The average elapsed times of all tested algorithms against the variable cardinality

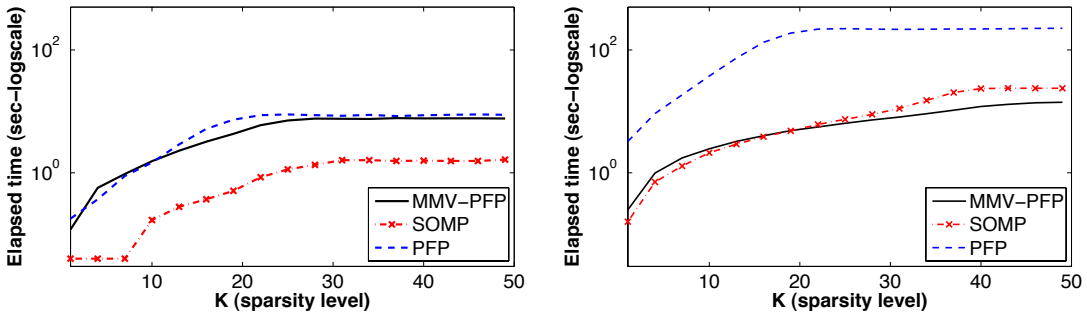


Figure 4.12: Average elapsed time (over 100 trials) for SOMP, MMV-PFP and PFP vs sparsity level K for a dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M = 50$, $N = 100$ and number of channels (a) $L = 2$ and (b) $L = 50$.

of the support set are depicted in Fig. 4.12(a)-(b). In the first experimental setting for which the number of multiple vectors was set at $L = 2$, it can be noted that SOMP achieves the fastest convergence for all sparsity levels examined. Regarding the PFP algorithms, MMV-PFP does not lead to any significant computational savings in this case, but as shown in Fig. 4.12(b) its convergence speed remains almost unaffected as the number of vectors increases, which results in MMV-PFP being the fastest algorithm for sparsity levels $K > 20$.

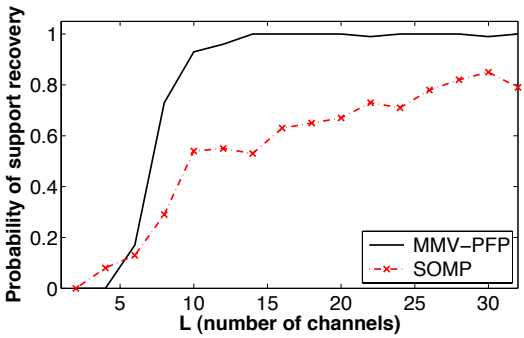


Figure 4.13: Probability of support recovery (over 100 trials) of SOMP and MMV-PFP vs the number of channels L for a dictionary $\mathbf{A} \in \mathbb{R}^{M \times N}$ with $M = 32$, $N = 256$ and fixed sparsity level at $K = 16$.

Finally, the recovery rates of MMV-PFP and SOMP algorithms were compared for a variable number of measurement vectors L, in the case when the sparsity level is fixed at $K = 16$ and the overcomplete dictionary $\mathbf{A} \in \mathbb{R}^{32 \times 256}$ with its column drawn from the normal distribution. The number of channels varied from $L = 2$ to $L = 32$ in steps of 2. In Fig. 4.13 the recovery results are depicted. As can be seen, although the performance of SOMP rises rapidly for small values of L, it becomes apparent that this

increase in performance slows down as L gets larger and it never achieves perfect recovery probability. On the other hand, the proposed method fails to recover the sparse matrix $\mathbf{X}$ when $L \leq 4$, but its performance displays a rapid increase for $4 < L \leq 12$ and stalls at almost perfect recovery for values of L larger than 14.

4.7 Discussion

This chapter investigated the applicability of structured sparsity, such as block structure and joint sparsity in the MMV setting, using the method of Polytope Faces Pursuit.

First of all, in the group sparsity case it was shown that the dual to the initial constrained least squares problem exists and the necessary and sufficient conditions for primal-dual feasibility were provided. Next, the Group Polytope Faces Pursuit algorithm was derived. The specific algorithm in analogy with the standard PFP algorithm, is a path following approach that exploits the geometry of the dual group sparsity problem. As described, of critical importance in the group sparse recovery problem is the choice of the objective function to be minimized and for this reason, both problems of $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization were examined. In spite of the fact that the resulting algorithms are suboptimal in the sense that they cannot ensure that the KKT conditions will always be met, it was shown that stronger theoretical guarantees exist for the identification of the true support of the unknown group sparse vector and hence its perfect recovery. These theoretical results differ for the two problems considered, namely $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization and they do not reveal which algorithm is expected to perform better in practice. However, the corresponding phase diagrams showed that although both algorithms enhance the recovery performance of the original PFP algorithm, GPFP based on $\ell_{2,1}$ minimization achieves the best phase transition. Further experimental results verified its superiority in recovery performance, while they also showed performance gains in certain cases against the popular GOMP algorithm.

Finally, the MMV problem under the assumption of joint sparsity was considered. Following a similar approach, the dual problem was derived and based on its geometry, PFP was extended resulting in the development of the MMV-PFP algorithm. The theoretical analysis presented did not manage to capture the expected performance gains over the conventional sparsity approach, resulting in rather worse theoretical results for

MMV-PFP as opposed to PFP. However, simulations on synthetic data demonstrated its superiority for the problem of joint sparse recovery. This suggests that an average case analysis that follows a probabilistic approach could give more insight on the performance gains of the MMV-PFP algorithm over the conventional sparsity approach of PFP.

The following chapters are concerned with the application of the developed sparse recovery algorithms to the classic problem of DOA estimation. From one point of view, this will allow the benchmarking of the performance of the proposed methods in a practical signal processing application, providing in that way new insights on the behaviour of the algorithms, especially in the noisy case. On the other hand, the developed methods are expected to provide certain improvements to the problem of DOA estimation and source localization.

Chapter 5

DOA estimation with sparsity constraints

As briefly described in Chapter 2, the problem of source localization and direction-of-arrival estimation appears in a large variety of signal processing applications and for this reason it has gained much interest among the academic society. Apart from the existing array processing methods for DOA estimation [13], the field of compressed sensing and sparse representations has played a crucial role in bringing a new perspective and giving more insight to the general problem of source localization. This impact has been evident throughout the recent years and cannot be overlooked when considering the overgrowing number of publications that attempt to exploit the compressed sensing framework by directly applying it to the problem of DOA estimation [122, 123, 124, 125, 126, 127] or by using its recent theoretical advancements in an oblique way as a tool for further analysis of the problem [128, 129, 130].

Among the aims of this PhD thesis is to develop efficient techniques based on sparse representations with application to the problem of DOA estimation. In previous chapters, the discussion was mainly concentrated around the more general topic of sparsity. After pointing out the main principles of the problem, efficient algorithms in terms of speed convergence and recovery performance were developed. The main objective was to provide robust algorithms designed to be applicable in real world applications and be able to exploit certain properties of the signals of interest. The focal point of the current chapter is to show how these methods and algorithms can be applied to the challenging problem of source localization.

For this reason, the problem of DOA estimation is initially formulated as a sparse recovery problem. In Section 5.3 an analysis of the theoretical properties of the problem is attempted following a compressed sensing approach. In that way, the advantages and limitations of this approach are identified, which eventually leads to the proposal of methods that show certain improvements over the existing sparsity based approaches.

5.1 Far-field DOA estimation with sparsity constraints

The discussion in this section begins with the problem of DOA estimation in the scenario of far-field propagation assuming narrowband sources impinging on a linear array sensors.

5.1.1 Problem formulation

Consider a uniform linear array (ULA) of M sensors with inter-element spacing d (Fig. 2.1). The sensors of the ULA are assumed to be calibrated. For simplicity and without loss of generality, it is also assumed that K plane waves propagating from the far-field impinge on the array from the unknown angles $\theta_1, \theta_2, \ldots, \theta_K$. Assuming no multi-path propagation and that the received signals on the array of sensors are narrowband with central frequency f_c, the m-th sensor captures a superimposition of the incoming signals with time delays (phase differences) of τ_m, which are functions of the signals' DOAs θ_i.

Therefore, taking as a reference sensor the one in the middle of the ULA, the linear array response to the impinging plane wave can be expressed as:

$$\mathbf{a}(\theta_i) = \left[e^{-j\omega_c\tau_1}, e^{-j\omega_c\tau_2}, \ldots, e^{-j\omega_c\tau_M}\right]^T \tag{5.1}$$

where ω_c is the angular frequency, $\tau_m = \left(m - \frac{M+1}{2}\right)d\cos(\theta_i)/c$ is the time delay and c is the speed of the propagation. The sensor spacing is chosen to be at half the wavelength i.e. $d = \lambda/2$ so that no spatial aliasing occurs. Considering that the wavelength of the plane wave is given by $\lambda = c/f_c$, the time delay at the m-th sensor can be expressed as $\tau_m = \frac{1}{2}\left(m - \frac{M+1}{2}\right)\cos(\theta_i)/f_c$. The substitution of the above expression for τ_m and $\omega_c = 2\pi f_c$ to equation (5.1) yields:

$$\mathbf{a}(\theta_i) = \left[e^{j\pi\frac{M-1}{2}\cos(\theta_i)},\ e^{j\pi\frac{M-3}{2}\cos(\theta_i)},\ \ldots,\ e^{-j\pi\frac{M-1}{2}\cos(\theta_i)}\right]^T. \tag{5.2}$$

Following the concept of spatial sparsity, as introduced by Malioutov et al. [125], the angular space is discretized into N possible angles of arrival so that $N > M$ and $N >> K$. The resulting overcomplete dictionary $\mathbf{A} \in \mathbb{C}^{M \times N}$ contains atoms corresponding to the impulse responses of the array for all potential N angles of arrival, such that:

$$\mathbf{A} = \big[\mathbf{a}(\theta_1),\, \mathbf{a}(\theta_2),\, \ldots,\, \mathbf{a}(\theta_N)\big]. \tag{5.3}$$

It follows that the sensor measurements at the ULA can be modelled as:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{n}(t) \tag{5.4}$$

where $\mathbf{x}(t)$ is a sparse vector containing K nonzero entries corresponding to the unknown DOAs, while $\mathbf{n}(t)$ is the additive noise at the sensors. In principle one could apply any sparse recovery algorithm i.e. ℓ_1 minimization to recover the unknown sparse vector $\mathbf{x}(t)$. However, in order to successfully estimate the parameters $\theta_1, \theta_2, \ldots, \theta_K$ the exact support recovery of $\mathbf{x}(t)$ is required, namely $\text{supp}(\hat{\mathbf{x}}(t)) = \text{supp}(\mathbf{x}(t))$ where $\hat{\mathbf{x}}(t)$ is the estimated sparse vector at the time instance t. It follows that any false detection on the estimated support set due to wrong atom selections, will introduce error in the parameter estimation problem. In other words, the problem of DOA estimation under the sparsity recovery framework should be viewed as a support recovery problem rather than a sparse approximation problem. This observation classifies the DOA estimation problem as a harder problem to the already hard task of sparse approximation.

5.1.2 Existing approaches

Nevertheless, instead of attempting to solve the inverse system of equation (5.4) for each time instant separately, assuming spatio-stationarity, namely that the unknown parameters do not vary with time but they rather remain constant and share a common sparse support of at most K indices, one could exploit the underlying joint sparsity and solve the inverse problem for several time samples simultaneously. Therefore, as explained in Section 4.5, the resulting MMV sparse recovery problem is formulated as:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{2,1} \quad \text{such that} \quad \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_{\text{F}} \leq \epsilon \tag{5.5}$$

where the $\mathbf{Y} \in \mathbb{C}^{M \times T}$ is the matrix containing sensor measurements of T time snapshots and $\mathbf{X} \in \mathbb{C}^{N \times T}$ is the jointly sparse vector. Comparing the problem of equation (5.5) with the problem (4.57) introduced in Section 4.5, one can notice that the constraint equality of (4.57) has been replaced by an inequality. This is to account for the additive noise in the sensor measurements and hence (5.5) illustrates the noisy mixed $\ell_{2,1}$ minimization problem. Even though the MMV formulation can be proven to provide better recovery performance and be less computational demanding than solving the SMV problem at each time instant, large numbers of snapshots T can increase dramatically the overall computational cost.

For that reason, the so called ℓ_1-SVD method, proposed in [3], includes a pre-processing step that reduces the dimensionality of the measurement matrix $\mathbf{Y}$ using the method of singular value decomposition (SVD):

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{5.6}$$

where $\mathbf{U}$ is an $M \times M$ unitary matrix, $\mathbf{\Sigma}$ is a diagonal matrix of size $M \times N$ containing the singular values and $\mathbf{V}^H$ denotes the Hermitian transpose of the unitary $N \times N$ matrix $\mathbf{V}$. In moderate noise levels, the singular values can provide information about the signal and the noise subspace. In contrast to the MUSIC algorithm that is merely concerned with the extraction of the noise subspace, the purpose of ℓ_1-SVD is to discard the noise subspace and keep only the K largest eigenvalues corresponding to the eigenvectors associated with the signal subspace. As explained in more detail in [3], this dimensionality reduction of the measurements is achieved by replacing the measurements matrix $\mathbf{Y}$ with $\mathbf{Y}_{\text{SV}} = \mathbf{Y}\mathbf{V}_K$, where $\mathbf{V}_K$ denotes the truncated matrix, obtained by keeping the K first columns of $\mathbf{V}$ and therefore $\mathbf{Y}_{\text{SV}} \in \mathbb{C}^{M \times K}$. It follows that the problem of equation (5.5) can be replaced by the following:

$$\min_{\mathbf{X}_{\text{SV}}} \|\mathbf{X}_{\text{SV}}\|_{2,1} \quad \text{such that} \quad \|\mathbf{Y}_{\text{SV}} - \mathbf{A}\mathbf{X}_{\text{SV}}\|_{\text{F}} \leq \epsilon' \tag{5.7}$$

where $\mathbf{X}_{\text{SV}}$ is an $N \times K$ matrix that relates to $\mathbf{X}$ in a similar way to how $\mathbf{Y}_{\text{SV}}$ relates to the measurements matrix $\mathbf{Y}$.

However, the problem of support recovery in the joint sparsity scenario is no longer

combinatorial when the number of given snapshots is larger than the support set [118]. If for the number of sources it holds that $K < T$ then the unknown support set can be identified by the powerful MUSIC algorithm. Therefore, considering the fact that in (5.7) SVD was used in order to reduce the dimension to K snapshots, one can argue that MUSIC could be used instead of mixed $\ell_{2,1}$ minimization. Nevertheless, the scope of the work in [118] is limited when it comes to the DOA estimation problem, as the analysis and the theoretical results presented are only concerned with the noiseless case. On the other hand, in the rank-deficient case i.e. when the sources impinging on the array are highly correlated, MUSIC is known to fail to resolve them and its DOA estimation performance degrades significantly. Moreover, its performance is also known to deteriorate in the presence of strong noise or when the number of snapshots taken are not sufficient for the estimation of the covariance matrix.

Fig. 5.1(a)-(b) show the spatial power spectrum for two independent sources impinging on a ULA of $M = 8$ sensors from the angles 60° and 72°. The noise level is at 15 dB and the available number of snapshots is $T = 5$. As can be seen, although both MUSIC and MVDR beamforming perform better than the classic beamforming method with regards to separating the sources, their peaks do not fall on the true DOAs and hence they both introduce error. However, the method based on the joint sparsity framework - which uses a dictionary of $N = 180$ atoms resulting from the uniform discretization of the angular space with resolution 1° - is able to perfectly recover the unknown support showing clear peaks at the true DOAs. On the other hand, in Fig. 5.1(c)-(d), which illustrate the analogous DOA estimation results in the case where the second source is located at 68°, ℓ_1-SVD fails to perfectly recover both DOAs. However, it still shows its superior performance over all other methods, since it depicts two clear peaks.

Because of the above, the following sections of the current chapter attempt to clarify the trade-off between array processing subspace based approaches and sparsity based methods and quantify their recovery performance and their limitations. Regarding the MMV sparse recovery formulation of equation (5.7), a further attempt is made in order to reduce its computational complexity by applying greedy algorithms that promote joint sparsity, such as MMV-PFP.

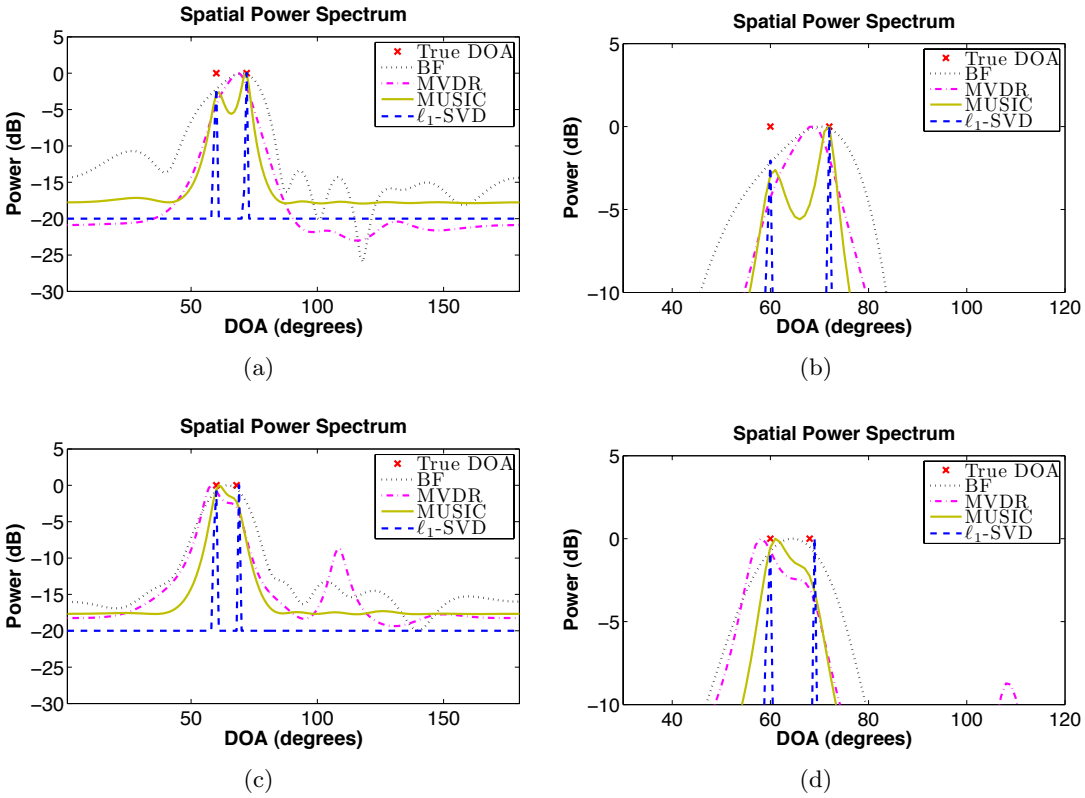


Figure 5.1: Spatial power estimate using conventional beamforming, MVDR, MUSIC and ℓ_1-SVD algorithms for two closely spaced sources with true DOAs (a)-(b) 60° and 72° and (c)-(d) 60° and 68°.

5.2 DOA dictionary coherence & discretization of the angular grid

The two most critical points for robust support recovery in compressed sensing are sparsity and incoherence. Therefore, regarding the DOA estimation problem, as long as the spatial sparsity assumption holds, the redundant dictionary $\mathbf{A}$ should exhibit low maximal coherence among its atoms to enable sparse recovery of the unknown support set with cardinality equal to the number of impinging sources. Although, in the specific setting the design of the dictionary is out of one's control due to the fact that its atoms are associated with the physics of the problem, there are still important parameters such as the redundancy of the dictionary given by the ratio M/N, the sensor spacing d and the discretization of the angular space that contribute to the dictionary mutual coherence and need to be selected with caution.

This section addresses the way of selecting the N-dimensional grid associated with the discretization of the angular space and shows how it affects the dictionary incoherence.

First of all, note that the coherence of the parametric dictionary $\mathbf{A}$ is given by:

$$\begin{aligned}
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{|\mathbf{a}^H(\theta_j)\mathbf{a}(\theta_i)|}{\|\mathbf{a}(\theta_j)\|\|\mathbf{a}(\theta_i)\|} \Rightarrow \\
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{1}{M}|\mathbf{a}^H(\theta_j)\mathbf{a}(\theta_i)| \Rightarrow \\
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{1}{M}\left|\sum_{m=1}^{M} e^{j\pi(m-\frac{M+1}{2})\cos\theta_j} e^{-j\pi(m-\frac{M+1}{2})\cos\theta_i}\right| \Rightarrow \\
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{1}{M}\left|\sum_{m=1}^{M} e^{-j\pi(m-\frac{M+1}{2})(\cos\theta_i-\cos\theta_j)}\right| \Rightarrow \\
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{\sin\big(M\pi(\cos\theta_i-\cos\theta_j)/2\big)}{M\sin\big(\pi(\cos\theta_i-\cos\theta_j)/2\big)} \Rightarrow \\
\mu &= \max_{1\leq i,i\neq j\leq N} \frac{\sin(M\pi\Delta_{i,j}/2)}{M\sin(\pi\Delta_{i,j}/2)}
\end{aligned} \tag{5.8}$$

where the exponential sum formula has been used. Equation (5.8) shows that the nearer the angles θ_i and θ_j are, the higher the coherence between them will be. In other words, for a given grid of N discrete locations, a dictionary atom $\mathbf{a}(\theta_i)$ will likely exhibit the highest correlation with one of each adjacent atoms i.e. $\mathbf{a}(\theta_{i-1})$ or $\mathbf{a}(\theta_{i+1})$. Furthermore, in the case that the grid is selected by uniformly discretizing the angular space, the overall dictionary coherence μ is expected to occur for two adjacent atoms such that the quantity $\Delta_{i,i+1} = \cos\theta_i - \cos\theta_{i+1}$ is minimized. Due to the nature of the cosine function, it follows that adjacent angles close to $\pi/2$ will show lower correlations than adjacent angles around 0 or π. Therefore, as also shown in [18], a better way to obtain the grid points is to uniformly discretize the cosine function instead of the angle θ. In that case, the quantity Δ will be constant for all adjacent angles and subsequently the overall dictionary coherence will be lower than the one resulting from the uniform discretization of the angular space.

To show this, note that in the case of uniformly discretizing the u-space ($u = \cos(\theta)$), the quantity Δ for any two adjacent grid points is constant and equal to $2/N$. Substituting this in equation (5.8), the coherence becomes:

$$\mu = \frac{\sin(M\pi/N)}{M\sin(\pi/N)}. \tag{5.9}$$

Therefore, according to (5.9) the dictionary coherence depends only on the redundancy

ratio M/N. As can be seen, the larger the value of M/N ratio, the lower the coherence μ. Note also that for large enough values of N, the approximation $\sin(\pi/N) \approx \pi/N$ holds. Therefore, equation 5.9 can be well approximated by the sinc function:

$$\mu \approx \frac{\sin(M\pi/N)}{M\pi/N} = \text{sinc}(M/N). \tag{5.10}$$

On the other hand, in the case of uniformly discretizing the angular space, for two adjacent grid points it holds that:

$$\Delta_{i,i+1} = \cos\theta_i - \cos\theta_{i+1} = \cos\theta_i - \cos(\theta_i + \pi/N). \tag{5.11}$$

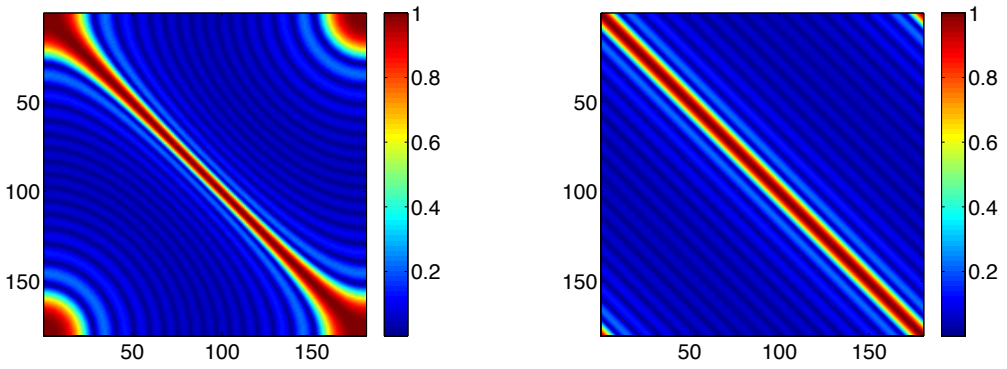
It is evident from equation (5.11) that the minimum value of $\Delta_{i,i+1}$, which results in the maximum dictionary coherence, is achieved at $\theta_i = k\pi$. Therefore, after the appropriate manipulations it can be shown that μ is given by:

$$\mu = \frac{\sin\Big(M\pi\big(1-\cos(\pi/N)\big)/2\Big)}{M\sin\Big(\pi\big(1-\cos(\pi/N)\big)/2\Big)} \tag{5.12}$$

where the angle sum trigonometric identity has been used. It follows that for values of $N >> \pi$ the approximation $\sin\Big(M\pi\big(1-\cos(\pi/N)\big)/2\Big) \approx M\sin\Big(\pi\big(1-\cos(\pi/N)\big)/2\Big)$ holds and hence $\mu \approx 1$. This means that the uniform discretization of the angular space will introduce ambiguities due to the near collinearities between atoms corresponding to the endfire angles, regardless of the redundancy ratio M/N. This could make it impossible for the CS recovery algorithms to distinguish between the endfire DOAs and therefore their performance is expected to considerably degrade. On the other hand, if the sources arrive from the broadside angles then they will be much easier resolved, since it holds:

$$\cos(\pi/2) - \cos(\pi/2 + \pi/N) = \sin(\pi/N) > \sin(2/N) \approx 2/N. \tag{5.13}$$

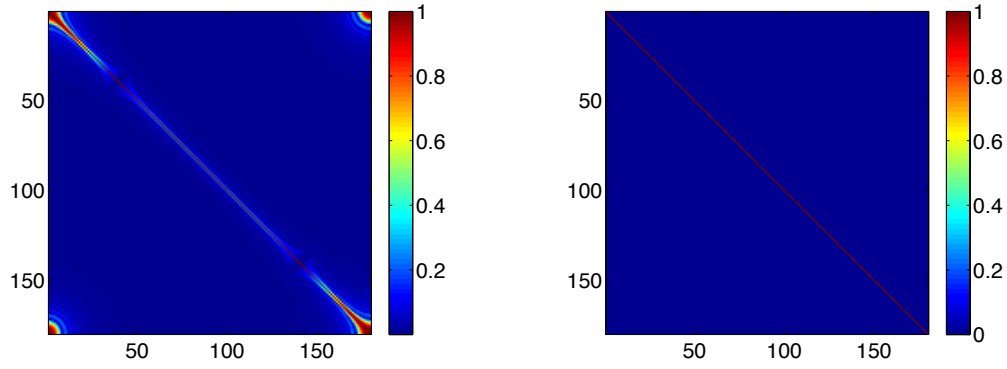
This can be also verified by considering the Gram matrix $\mathbf{G} = |\mathbf{A}^H\mathbf{A}|$ for both discretization scenarios. Fig. 5.2 displays as images the gram matrices $\mathbf{G}$ in the case of uniformly discretizing the DOA space (Fig. 5.2(a)) and in the case of uniformly discretizing the u-space (Fig. 5.2(b)).



(a) (b)

Figure 5.2: Gram matrix of a DOA dictionary corresponding to a ULA of $M = 8$ sensors when the grid points are obtained by uniformly discretizing (a) the angular space (b) the u-space.

It is also worth mentioning that in the latter case, if the number of sensors is equal to the number of grid points $M = N$, then the resulting dictionary is a basis i.e. FFT basis. Hence, as can be seen in Fig. 5.3(b) $\mathbf{A}$ exhibits maximal incoherence. However, the same does not hold in the case of discretizing the angular space (Fig. 5.3(a)).



(a) (b)

Figure 5.3: Gram matrix of a square DOA dictionary ($M = N$) corresponding to a ULA when the grid points are obtained by uniformly discretizing (a) the angular space (b) the u-space.

In what follows, dictionaries obtained by uniformly discretizing of the u-space are considered, due to their better overall properties.

5.3 Geometry of the array (spacing and number of sensors)

Consider now the RIP property of (2.49). It is required that $\delta_{2K} < 1$ for the solution to the combinatorial problem (4.56) to be unique. Given a dictionary $\mathbf{A} \in \mathbb{C}^{M \times N}$ and any subset S of its columns with cardinality K, it also holds:

$$\delta_{2K} = \max_{S} \|\mathbf{A}_{2S}^{H}\mathbf{A}_{2S} - \mathbf{I}\|_2 \leq \max_{S} \|\mathbf{A}_{2S}^{H}\mathbf{A}_{2S} - \mathbf{I}\|_1 \leq (2K-1)\mu \tag{5.14}$$

where the first inequality is a result of the fact that the matrix $\mathbf{A}_{2S}^{H}\mathbf{A}_{2S} - \mathbf{I}$ is Hermitian [55]. Therefore, it is required that $(2K-1)\mu < 1$ or after substituting equation (5.10) $\text{sinc}(M/N) < 1/(2K-1)$. Although, the last inequality cannot be solved algebraically for M without using approximations such as the Taylor series, one can obtain the value for M graphically, finding the intersection between $\text{sinc}(M/N)$, which is a function of M for fixed N and the line $1/(2K-1)$ for a specified sparsity level K. As can be seen in Fig. 5.4, for a grid of $N = 180$ points and sparsity level $K = 3$ the number of sensors should be greater than $M = 148$ for perfect support recovery.

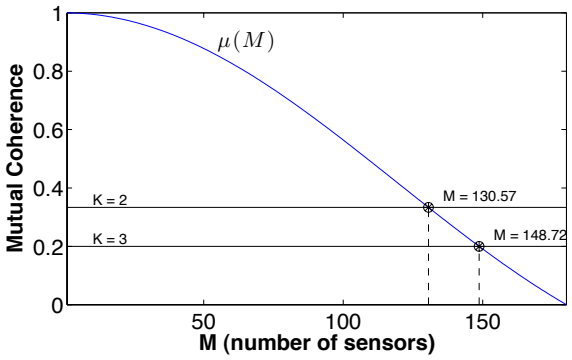


Figure 5.4: Mutual coherence of the DOA dictionary as a function of the number of sensors M.

Note that the resulting bound on M using the above analysis is approximate, since one can precisely identify M by solving $\max_S \|\mathbf{A}_{2S}^{H}\mathbf{A}_{2S} - \mathbf{I}\|_2 < 1$. This, however, is a much harder problem, since it involves finding the largest maximum eigenvalue of $\mathbf{A}_{2S}^{H}\mathbf{A}_{2S} - \mathbf{I}$ among all possible combinations of atoms of cardinality $2K$.

In any case, the resulting lower bound on M based on the RIP is quite pessimistic. As it will be shown in later sections, even with a much smaller number of sensors M one can achieve super-resolution, as long as the assumption that the location of the impinging sources exactly match the ones included in the angular grid and the noise level

is moderate. Besides, RIP does not take into account the additional information provided by the multiple time snapshots, which allow the formulation of the DOA estimation problem as a joint sparse recovery problem. As explained in Section 4.5.2 of the previous chapter, although worst case theoretical guarantees fail to predict any performance gain, empirical findings demonstrated that the MMV model can boost the performance of the conventional sparsity approach based on the SMV model.

Nevertheless, the recovery capability of the proposed method could still be further improved by considering random non-uniform linear arrays (NLA) (Fig. 5.5), which however implies longer aperture arrays.

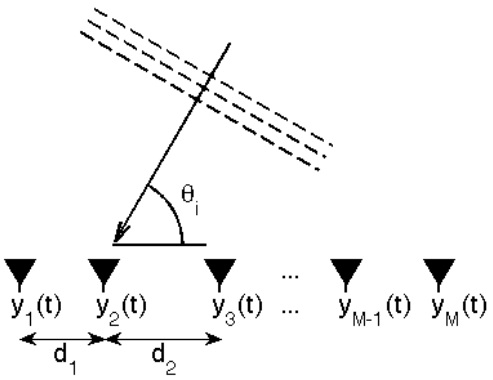


Figure 5.5: Impinging plane wave on a non-uniform linear array of M sensors.

As mentioned previously and also explained in [131], the complete DOA dictionary $\mathbf{A}$ resembles the FFT ensemble. To see this, consider once again the function $\mathbf{a}_m(\theta_i) = \exp\Big\{ -j\pi\big(m - \frac{M+1}{2}\big)\cos\theta_i\Big\}$. The u-space is uniformly discretized into N grid points, where $u = \cos\theta$ and $u \in [-1, 1]$. Hence, each angle of the grid is given by:

$$\theta_i = \cos^{-1}\left(\frac{2i - N}{N}\right). \tag{5.15}$$

Using now (5.15), the linear array response becomes $\mathbf{a}_m(\theta_i) = \exp\Big\{-j\pi\big(m-\frac{M+1}{2}\big)\frac{2i-N}{N}\Big\}$, and after further manipulation, it can be shown that $\mathbf{a}_m(\theta_i)$ is given by:

$$\mathbf{a}_m(\theta_i) = \exp\Big\{ -j\frac{2\pi}{N} d_m c_i\Big\} \quad \text{for} \quad 1 \le m \le M, \quad 0 \le i \le N-1 \tag{5.16}$$

where $d_m = m - \frac{M+1}{2}$, $c_i = i - N/2$. Setting $M = N$, the resulting DOA dictionary is a Fourier matrix and subsequently, results from the field of compressed sensing can be used to theoretically demonstrate the advantages of using an NLA over a ULA. More specifically, considering the complete square matrix and randomly selecting M of its N rows, the theoretical results in [72] for the uniformly random partial Fourier matrix will also hold for the NLA DOA dictionary. Therefore, the RIP will be satisfied with high probability as long as $M \approx Ck\log^4(N)$.

The above result has also been improved in [132], where the authors consider a weak-RIP analysis, which stands as a weaker alternative to the RIP condition and more practical for applications [133]. A matrix $\mathbf{A} \in \mathbb{K}^{M\times N}$ is said to satisfy the weak-RIP if:

$$(1-\delta_{K+t})\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}_J\mathbf{x}\|_2^2 \leq (1+\delta_{K+t})\|\mathbf{x}\|_2^2 \quad \forall J \supset S \tag{5.17}$$

where $|J| = K+t$ and δ_{K+t} denotes the corresponding weak restricted isometry constant. The work in [132] considers the case with $t = 1$ and shows that a uniformly random partial Fourier matrix satisfies the weak-RIP for any ϵ,$\delta \in (0,1)$, if:

$$M \geq \left(\frac{2(3+\delta)}{3\delta^2}\right)\left\{\log\left(\frac{2(N-K)}{\epsilon}\right) + \log(K+1)\right\}(K+1) \tag{5.18}$$

with probability $\mathbb{P}(\delta_{K+1} \geq \delta) \leq \epsilon$.

These results suggest that an improved recovery performance should be expected when employing NLA for the DOA estimation problem for the same number of sensors M or that the same performance with a ULA array can be achieved using fewer sensors.

5.4 MMV algorithms for DOA estimation

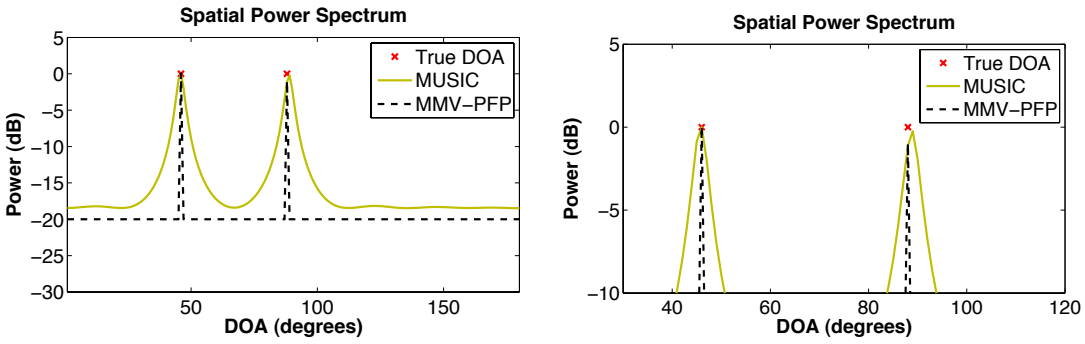
It has already been mentioned that the sparse representations framework was first introduced in the field of DOA estimation and source localization in the pioneering work in [3]. The specific work mainly focuses on uniform linear arrays and proposes uniform discretization of the angular space. Therefore, it is expected that the performance of the ℓ_1-SVD method will improve considerably by taking into account the results of the previous sections.

However, although the SVD decomposition of the multiple measurement vectors $\mathbf{Y}$

can considerably reduce the complexity of the optimization task, still solving the problem of equation with SOCP methods can be computationally demanding in practical applications. To this end, the specific work focuses on the applicability of faster alternative greedy approaches (i.e. SOMP, MMV-PFP etc.).

5.4.1 Well separated sources

Aside from reducing the overall complexity of the mixed norm minimization problem using a greedy MMV algorithm, when the number K of sources impinging on the array is known a priori, this information can be easily incorporated by forcing the greedy algorithm to terminate after only K iterations. This can be very beneficial especially in the cases where the variance of the additive noise is unknown. Additionally, it could further reduce the overall computational cost. Fig. 5.6 shows the estimated spatial power spectrum using MUSIC and MMV-PFP for a given ULA of $M = 8$ sensors. The size of the angular grid is selected to be $N = 180$ grid points. The number of sources is $K = 2$ and the corresponding plane waves are assumed to arrive from the angles 46° and 88°. As can be seen, although both algorithms show two clear peaks, MMV-PFP exactly resolves the spatial spectrum resulting in accurate localization of both sources (Fig. 5.6(b)).



(a) (b)

Figure 5.6: Spatial power estimate of MUSIC and MMV-PFP algorithms for two well separated sources with true DOAs 46° and 88°.

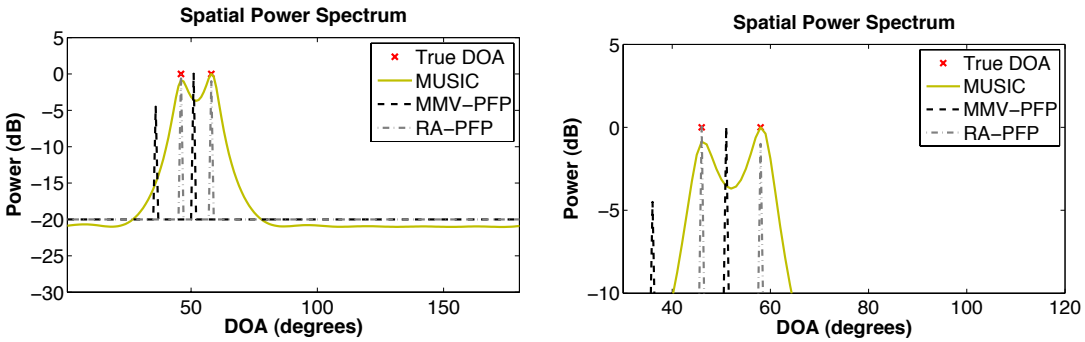
5.4.2 Closely spaced sources

However, the proposed MMV-PFP method might fail to resolve closely spaced sources. To overcome this limitation, inspired by the work in [118], where the authors develop rank aware joint sparse recovery methods, a similar modification of the MMV-PFP algorithm is proposed to aid the identification of the support set when the dictionary coherence is high.

More specifically, after each iteration of the MMV-PFP algorithm, a sensing dictionary $\tilde{\mathbf{A}}$ is constructed by orthogonalizing the set of atoms that have not yet been selected, in order to reduce the dictionary coherence. Then, the residual is also orthogonalized (i.e. using SVD decomposition) and projected to the sensing dictionary $\tilde{\mathbf{A}}$. The resulting Rank Aware MMV-Polytope Faces Pursuit (RA-PFP) algorithm is summarized in Algorithm 9.

Algorithm 9 Rank Aware MMV-Polytope Faces Pursuit

Input: $\mathbf{A}$, $\mathbf{Y}$
Set stopping conditions $l_{\max}$ and $\theta_{\min}$
Set $\tilde{\mathbf{A}} = \mathbf{A}$
Initialize:
$k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{C}^k \leftarrow \mathbf{0},\ \mathbf{X}^k \leftarrow \emptyset,\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{0},\ \mathbf{V}^k \leftarrow \emptyset\ \mathbf{R}^k \leftarrow \mathbf{Y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\max_i \mathbf{a}_i^H \mathbf{R}^{k-1} > \theta_{\min}$ **do** {Find next face}
  $k \leftarrow k + 1$
  Orthogonalize residual: $\mathbf{U}^k \leftarrow \text{orth}(\mathbf{R}^{k-1})$
  Find face:
  $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \{\|\tilde{\mathbf{a}}_i^H \mathbf{U}^k\|_2^2 / (\|\tilde{\mathbf{a}}_i^H \mathbf{U}^k\|_2 - \tilde{\mathbf{a}}_i^H \mathbf{C}^{k-1} (\mathbf{U}^k)^H \tilde{\mathbf{a}}_i)\}$
  Add constraints:
  $\mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{a}_{i^k}],\ \mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\}$
  $\mathbf{X}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{Y},\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{A}^k \mathbf{X}^k$
  $\mathbf{V}^k = [\mathbf{V}^{k-1}, \mathbf{a}_{i^k} \mathbf{R}^{k-1} / \|\mathbf{a}_{i^k} \mathbf{R}^{k-1}\|_2],\ \mathbf{C}^k \leftarrow [(\mathbf{A}^k)^\dagger]^H \mathbf{V}^k$
  Calculate orthogonal projection: $P_{\mathbf{A}^k}^\perp = \mathbf{I} - \mathbf{A}^k (\mathbf{A}^k)^\dagger$
  $\tilde{\mathbf{A}} \leftarrow P_{\mathbf{A}^k}^\perp \mathbf{A},\ \mathbf{R}^k \leftarrow P_{\mathbf{A}^k}^\perp \mathbf{Y}$
  Normalize: $\tilde{\mathbf{a}}_i \leftarrow \tilde{\mathbf{a}}_i / \|\tilde{\mathbf{a}}_i\|_2$
end while
Output: $\mathbf{X}^* = \mathbf{X}^k$



(a) (b)

Figure 5.7: Spatial power estimate of MUSIC, MMV-PFP and RA-PFP algorithms for two closely spaced sources with true DOAs 46° and 58°.

Fig. 5.7 shows the estimated spatial spectrum of all algorithms including the modified version of MMV-PFP, using the same setting with the one in Fig. 5.6 apart from the fact that the two incoming sources arrive from the angles 46° and 58°. RA-PFP shows its

advantage in identifying the DOAs of the closely spaced sources. In the following sections, more experimental results are presented in order to better quantify the performance of the proposed greedy algorithms.

5.5 Experimental results

In this section, further experimental results are presented and the performance of the MMV-PFP and RA-PFP algorithms is compared against sparsity based approaches such as ℓ_1-SVD and Rank Aware Order Recursive Matching Pursuit (RA-ORMP) [118] and the classical subspace based MUSIC algorithm. In all following experiments, it is assumed that the unknown DOAs fall into the predefined discrete set of angles. Both ULAs and NLAs are employed. Furthermore, it is also assumed that the number of sources, corresponding also to the spatial sparsity level K, is known a priori.

5.5.1 DOA estimation vs SNR

The first set of experiments attempts to evaluate the performance of DOA estimation for variable noise levels. For this reason, the specific experiment considers $K = 2$ narrowband sources impinging on a linear array of sensors from the far-field. The discrete set of potential DOAs was obtained by discretizing the u-space and it was assumed that the sources' true DOAs exactly match a subset of the N-dimensional discrete set of angles. N was fixed at 180 potential angles. The true DOAs were selected at random from the discrete set and the variance of noise level was varied resulting in SNR values ranging from -25 dB to 30 dB in steps of 5 dB. Fig. 5.8 depicts the DOA estimation performance, given as a probability of perfect support recovery vs the SNR. In particular, Fig. 5.8(a)-(b) illustrate the recovery probabilities of a ULA with $M = 8$ sensors for $T = 8$ and $T = 200$ time snapshots, while Fig. 5.8(c)-(d) show the corresponding results for an NLA. All results have been averaged over 100 trials.

It can be noted that in the case of a ULA the rank aware versions of MMV-PFP and SOMP yield the best overall performance. On the other hand, although ℓ_1-SVD and MMV-PFP algorithms perform close to the rank aware sparse methods for noise levels below 10 dB, for higher values of SNR their performance seems to stall at lower recovery rates compared to the rank aware algorithms. This might be due to the high mutual coherence of the DOA dictionary in this setting, resulting in decreased performance most

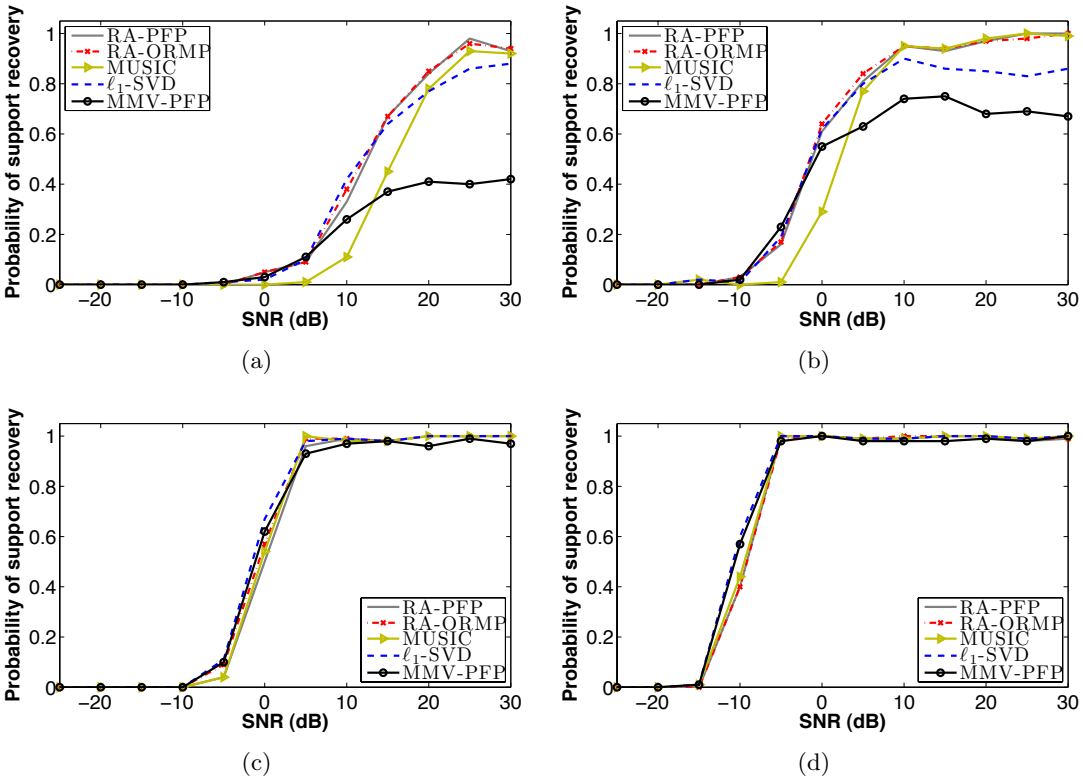


Figure 5.8: Probability of exact DOA estimation of $K = 2$ impinging on a ULA of $M = 8$ sensors for (a) $T = 8$ and (b) $T = 200$ samples and an NLA of $M = 8$ sensors for (c) $T = 8$ and (d) $T = 200$ vs SNR.

notably affecting MMV-PFP algorithm. Regarding the MUSIC algorithm, as expected it seems to be more prone to noise and as can be seen its support recovery performance degrades considerably when the SNR is less than 15 dB. In the case of $T = 200$ samples (Fig. 5.8(b)), DOA estimation performance improves for all tested algorithms, but similar trends are observed.

Fig. 5.8(c)-(d) show the results obtained for linear arrays of sensors with non-uniform spacing. In this case, all algorithms take advantage of the NLA structure resulting in improved recovery performance in both cases of $T = 8$ and $T = 200$ samples. Interestingly, all tested algorithms perform very close.

Next, the above experiment was repeated for an increased number of sensors $M = 16$. Fig. 5.9(a)-(d) summarize the obtained results that show improved recovery rates for all algorithms with similar trends to the case of $M = 8$ sensors.

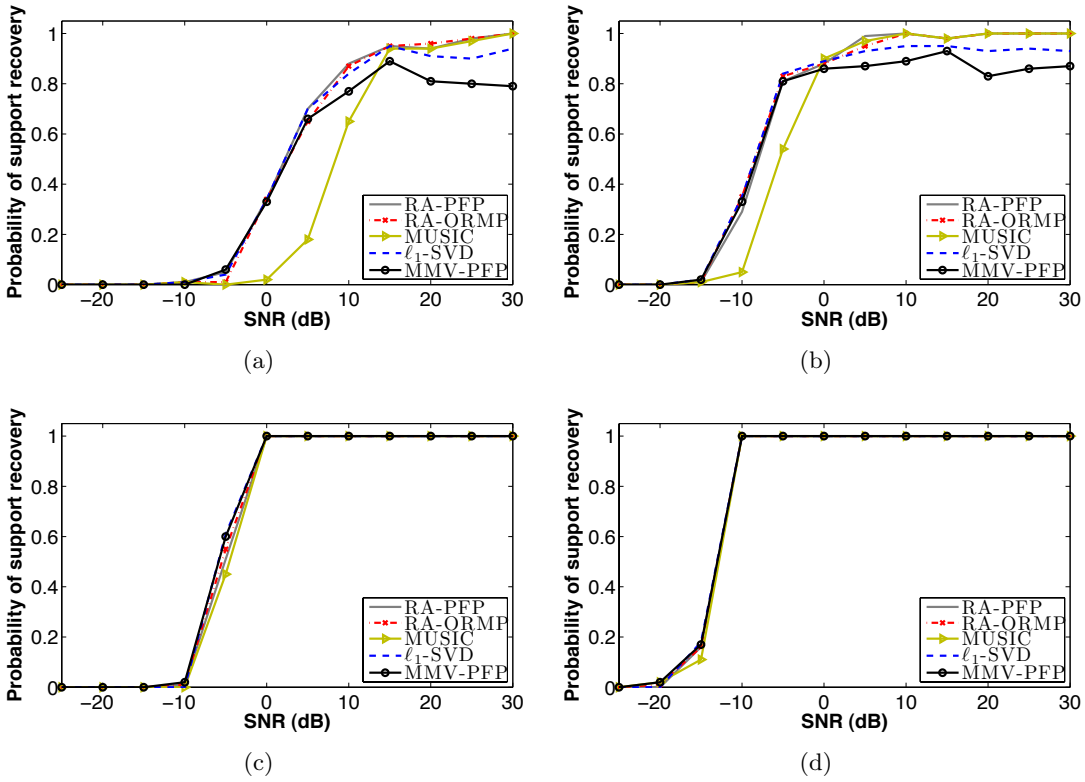


Figure 5.9: Probability of exact DOA estimation of $K = 2$ impinging on a ULA of $M = 16$ sensors for (a) $T = 8$ and (b) $T = 200$ samples and an NLA of $M = 16$ sensors for (c) $T = 8$ and (d) $T = 200$ vs SNR.

5.5.2 Number of sources

The second experiment tried to assess the DOA estimation performance of all examined algorithms as a function of the number of sources K, assuming that they arrive on the linear array from different angles.

To do this, a linear array of $M = 16$ sensors was considered and K synthetic complex-valued signals were generated at each instance with DOAs being uniformly at random selected from the predefined grid of $N = 180$ potential angles. The noise level was kept constant at SNR = 20dB. The experiment was initially run for a ULA and then repeated for NLA structures. The obtained results, depicted in Fig. 5.10, were the averages over 100 trials. It can be noted that in the case of a ULA and for $T = 8$ time snapshots (Fig. 5.10(a)) the proposed method of RA-PFP, followed by RA-ORMP, outperforms all other algorithms and provides a significant improvement over the original MMV-PFP algorithm, which yields the poorest recovery rates. When the number of snapshots is increased to $T = 200$ (Fig. 5.10(b)), all algorithms benefit, providing improved estimates.

In this specific setting, MUSIC, along with the rank aware joint sparsity algorithms, provides superior performance over ℓ_1-SVD and MMV-PFP.

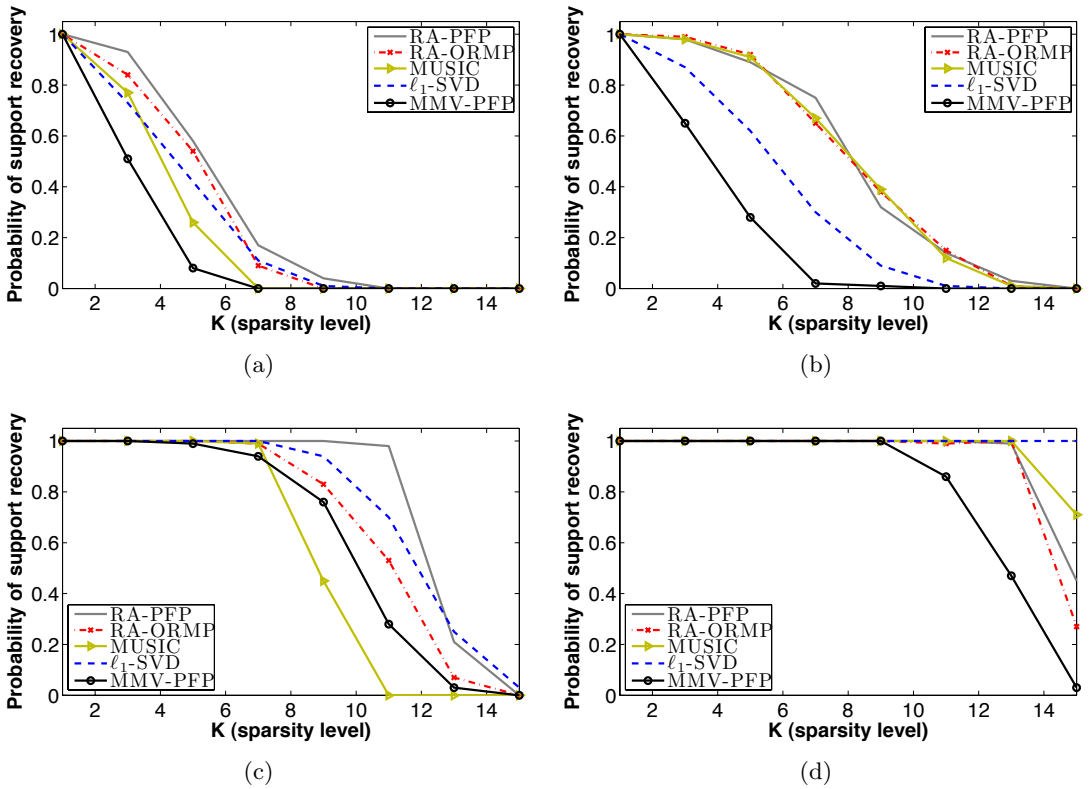


Figure 5.10: Probability of exact DOA estimation using a ULA of $M = 16$ sensors for (a) $T = 8$ and (b) $T = 200$ samples and an NLA of $M = 8$ sensors for (c) $T = 8$ and (d) $T = 200$ vs the number of sources K corresponding to the sparsity level. The SNR is fixed at 20 dB.

The support recovery performance improves significantly for NLA structures. As shown in Fig. 5.10(c) and in agreement with the theoretical analysis presented in Section 5.3, all algorithms improve their recovery ability. The proposed method of RA-PFP achieves the best overall performance succeeding in almost fully recovering the unknown DOAs for sparsity levels up to $K = 11$. Furthermore, the performance of MMV-PFP and ℓ_1-SVD is boosted in that case, with ℓ_1-SVD outperforming RA-ORMP and MMV-PFP. This shows once again the importance of the dictionary mutual incoherence in solving underdetermined inverse problems. Since using NLAs leads to dictionaries with lower maximal coherence, the signal subspace becomes more distinguishable and therefore the sparse recovery algorithms can easier retrieve the unknown support set of atoms. Taking additional time snapshots ($T = 200$) on top of that, can further enhance the performance of joint sparsity algorithms (Fig. 5.10(d)). In particular, ℓ_1-SVD achieves

exact estimation of the unknown directions for all cases examined, verifying the average case analysis result in 4.5.2, where it was shown that joint sparsity algorithms can recover supports with up to $M-1$ cardinality.

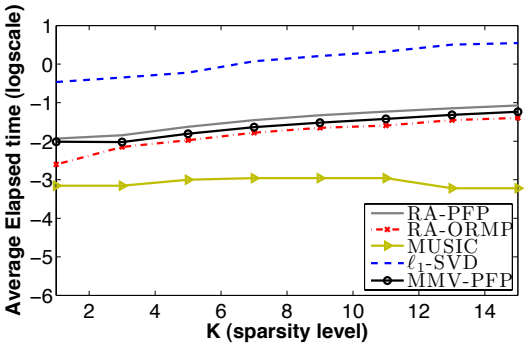


Figure 5.11: Average elapsed time of DOA estimation using a ULA of $M = 16$ sensors for $T = 200$ snapshots vs the number of sources K corresponding to the sparsity level.

Fig. 5.11 shows the average elapsed time for all algorithms considered in the case of a ULA of $M = 16$ and $T = 200$ time snapshots against the number of sources K. MUSIC exhibits the fastest convergence, which also remains almost constant with K. This is to be expected, since the algorithm involves an eigenvalue decomposition for the estimation of the noise subspace followed by a projection to the column space of the DOA dictionary $\mathbf{A}$. In contrast, the elapsed time of joint sparsity based algorithms is slightly affected by the increase of the sparsity level. As can be seen, for all sparse approaches, the average elapsed time rises almost linearly with a slight slope as the sparsity level K increases. Among them, RA-ORMP yields the fastest convergence followed closely by MMV-PFP and RA-PFP algorithms. When the sparsity level is $K = 5$, the average elapsed times of MUSIC, RA-ORMP, MMV-PFP, RA-PFP and ℓ_1-SVD are 1 msec, 10.6 msec, 15.6 msec, 23.7 msec and 599.8 msec respectively. Therefore, in that case the joint sparsity algorithms provide a significant speed gain over ℓ_1-SVD being at least 25 times faster. However, they are at the same time at least 10 times slower than MUSIC.

5.5.3 Spacing of the sources in the angular grid

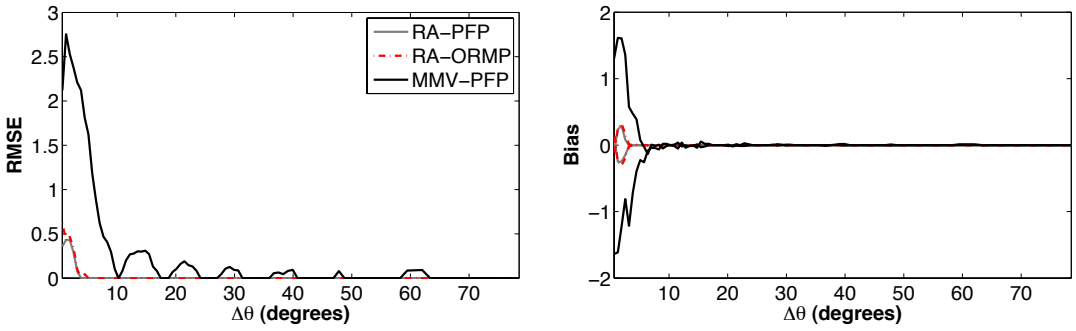
This section presents results that assess the resolution capability of the proposed method for DOA estimation. For the specific experiment, a ULA of $M = 16$ sensors was considered. After uniformly discretizing the u-space, a grid of $N = 180$ angles was obtained resulting in the overcomplete dictionary $\mathbf{A} \in \mathbb{C}^{16\times 180}$. The number of sources was fixed

at $K = 2$. Between the two impinging sources, the first was set at 90°, while the DOA of the second was left to vary from 90° to 180° with the step size equal to the minimum distance between two consecutive angles of the angular grid. Therefore, in this experiment the goal was to see how the sparse recovery algorithms performance is affected by the angular spacing of the sources $\Delta\theta$. To do this, for each value of $\Delta\theta$, the greedy joint sparsity algorithms was run for $n = 100$ trials and their recovery performance was measured using the root-mean-square error (RMSE) defined as:

$$RMSE(\hat{\boldsymbol{\theta}}) = \sqrt{\mathrm{E}\left\{\sum_{i=1}^{K}(\hat{\theta}_i - \theta_i)^2\right\}} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{K}\sum_{i=1}^{K}(\hat{\theta}_{ij} - \theta_i)^2} \tag{5.19}$$

and the bias function, which is given by:

$$Bias(\hat{\theta}_i) = \mathrm{E}\{\hat{\theta}_i - \theta_i\} = \frac{1}{n}\sum_{j=1}^{n}(\hat{\theta}_{ij} - \theta_i). \tag{5.20}$$



(a) (b)

Figure 5.12: (a) RMSE and (b) estimation bias as a function of the angular distance between $K = 2$ sources impinging on a ULA of $M = 16$ sensors.

Fig. 5.12(a) shows the RMSE as a function of the angular distance between the sources' DOAs. First of all, it can be observed that the rank aware joint sparsity algorithms behave much better in resolving closely spaced sources, as the estimation error remains at low levels even for the minimum separation $\Delta\theta$. For values of $\Delta\theta$ larger than 5° the recovery is always exact. On the other hand, for MMV-PFP, the error grows large for closely spaced sources and as can be seen even in the case of well separated sources the algorithm introduces some error occasionally. Although, this error is expected to vanish with the increase of the number of sensors, the presented experimental results reveal the

Table 5.1: Probability of exact support recovery in DOA estimation in the coherent source scenario.

			T = M/2	T = 50	T = 200
M = 16	SNR = 10dB	MUSIC	0.01	0.21	0.69
		RA-PFP	0.58	0.56	0.76
		MMV-PFP	**0.98**	**1**	**1**
		RA-ORMP	0.43	0.39	0.72
	SNR = 20dB	MUSIC	0.45	1	1
		RA-PFP	**0.65**	**1**	**1**
		MMV-PFP	**1**	**1**	**1**
		RA-ORMP	0.49	1	1

superiority of the rank aware algorithms. In Fig. 5.12(b), the bias of the DOA estimation is plotted for all algorithms. The bias shows in more detail the average distance of the estimated DOAs from the true DOAs for each source.

5.5.4 Correlated sources

The last experiment addresses the problem of DOA estimation when the sources impinging on the array are coherent or highly correlated. It is well known that the performance of MUSIC degrades significantly in this scenario as opposed to the ℓ_1-SVD method. This is one of its main advantages over the classic array processing methods [3]. In this experimental setting, once again a ULA of $M = 16$ sensors and an angular grid of $N = 180$ potential DOAs were considered. Firstly, two sequences of uncorrelated normal distributed random numbers $\mathbf{x}_1$ and $\mathbf{y}_1$ were generated. Next, a new sequence was generated by combining $\mathbf{x}_1$ and $\mathbf{y}_1$:

$$\mathbf{x}_2 = \rho\mathbf{x}_1 + \sqrt{1-\rho^2}\mathbf{y}_1 \tag{5.21}$$

so that the sequences $\mathbf{x}_1$ and $\mathbf{x}_2$ have a correlation equal to ρ. For this experiment ρ was set at the high value of 0.99. The Table 5.1 summarizes the resulting average (over 100 trials) support recovery results.

The obtained average recovery rates show that MMV-PFP is the least affected algorithm by the strong correlation between the incoming sources. Both rank aware algorithms' performance degrades although the proposed RA-PFP displays slightly better performance than RA-ORMP. This performance deterioration might be due to the fact that both algorithms incorporate the step of orthogonalization of the residual at each iteration, using SVD decompositon and QR factorization, respectively. Therefore, it seems

that the orthogonalization step tends to underestimate the signal subspace. However, when compared to MUSIC algorithm both methods perform considerably better.

5.6 Wideband DOA estimation

All previous sections of the current chapter considered narrowband linear arrays and assumed either that the propagating sources occupy narrow bands in the frequency spectrum or that they are broadband signals, which can be well approximated by digital modulation over a carrier frequency. However, when dealing with wideband sources i.e. speech or music signals, this narrowband approximation will no longer hold and consequently the obtained delayed signals cannot be represented by phase shifts.

One straightforward approach that has been extensively used for array processing methods [135, 136] such as MUSIC, stems from the Fourier transform convolution theorem, which states that a convolution in time is a multiplication in frequency and vice versa. In particular, according to this technique the measured data is initially transformed to the time-frequency domain. The time-frequency spectrum is then divided into multiple small regions, such that standard narrowband processing can be applied to each one of those separately. Therefore, after applying the Short-Time Fourier Transform (STFT), any DOA estimation narrowband algorithm can be utilized at each frequency bin individually. The final estimate is obtained by combining the individual estimates.

5.6.1 Far-field wideband DOA estimation

The far-field scenario, which allows the assumption of plane wave arrival is first considered. Suppose that a ULA of M sensors receives K sources propagating from K distinct angles and each sensor captures a superimposition of time delayed signals. The measurements at each sensor are partitioned into L time segments and each segment is transformed into Q frequency regions using the discrete Fourier transform. As long as the bandwidth of each subband is much smaller than its central frequency, corresponding to the frequency bin f_q, the narrowband assumption holds and it enables the following approximation [21, 137]:

$$y_m(t, f_q) = \sum_{i=1}^{K} e^{-j2\pi f_q (m - \frac{M+1}{2}) d \cos(\theta_i)/c} \mathbf{x}_i(t, f_q) + n_m(t, f_q). \tag{5.22}$$

Therefore, it becomes evident that one can apply the narrowband sparse representation DOA estimation framework at each frequency bin separately. After discretizing the angular space and forming the overcomplete dictionary $\mathbf{A}_q \in \mathbb{C}^{M\times N}$, one can attempt to solve the optimization problem:

$$\min_{\mathbf{X}_q} \|\mathbf{X}_q\|_{2,1} \quad \text{such that} \quad \|\mathbf{Y}_q - \mathbf{A}_q\mathbf{X}_q\|_{\mathrm{F}} \leq \epsilon \tag{5.23}$$

where it is assumed joint temporal sparsity over the unknown $\mathbf{X}_q$ and the subscript q indicates the q-th frequency bin. Furthermore, it is observed that the cardinality of the support of the joint sparse matrix $\mathbf{X}_q$ will be at most K at each frequency bin. Therefore, this underlying additional structure of joint sparsity in the frequency domain could be exploited by merging all different frequency bins that correspond to the same DOA and appropriately interleaving the entries of the measurement matrix $\mathbf{Y}_q$ and the unknown solution matrix $\mathbf{X}_q$. In that case, the dictionary is formed as:

$$\bar{\mathbf{A}} = \sum_{q=1}^{Q} \mathbf{A}_q \otimes \mathbf{I}_q \tag{5.24}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{I}_q \in \mathbb{C}^{QM\times QN}$ is a matrix with all elements zero but the q-th diagonal entry equal to one. The measurements matrix $\bar{\mathbf{Y}}$ can be obtained equivalently. The corresponding inverse optimization problem can then be formulated as:

$$\min_{\bar{\mathbf{X}}} \sum_{i=1}^{N} \|\bar{\mathbf{X}}_i\|_{2,2} \quad \text{such that} \quad \|\bar{\mathbf{Y}} - \bar{\mathbf{A}}\bar{\mathbf{X}}\|_{\mathrm{F}} \leq \epsilon. \tag{5.25}$$

The problem in 5.25 reveals additional group structure in the frequency domain as well as joint sparsity in the temporal domain. Therefore, the greedy MMV sparse recovery algorithms can be modified appropriately, in order to incorporate this additional structure.

5.6.2 A greedy algorithm for grouped frequency wideband DOA estimation

The problem of (5.25) is convex and its solution can be accomplished using standard software packages (i.e. SeduMi). However, in order to reduce the computational complexity, a simple greedy algorithm based on the combination of SOMP and GOMP algorithms is proposed. More specifically, the algorithm is identical to SOMP but uses a modified atom

selection criterion at each step in order to impose the group sparsity over all frequencies. The so-called Group Simultaneous Orthogonal Matching Pursuit (G-SOMP) is shown in Algorithm 10.

Algorithm 10 Group Simultaneous Orthogonal Matching Pursuit

Input: $\mathbf{A}$, $\mathbf{Y}$
Set stopping conditions $l_{\max}$ and $\epsilon_{\min}$
Initialize: $k \leftarrow 0,\ \mathcal{I}^k \leftarrow \emptyset,\ \mathbf{A}^k \leftarrow \emptyset,\ \mathbf{X}^k \leftarrow \emptyset,\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{0},\ \mathbf{R}^k \leftarrow \mathbf{Y}$
while $|\mathcal{I}^k| < l_{\max}$ and $\|\mathbf{R}^k\|_F > \epsilon_{\min}$ **do** {Find next group of atoms}
    $k \leftarrow k + 1$
    $i^k \leftarrow \arg\max_{i \notin \mathcal{I}^{k-1}} \|\mathbf{A}_i^H \mathbf{R}^{k-1}\|_{2,2}$
    Add constraints:
    $\mathcal{I}^k \leftarrow \mathcal{I}^{k-1} \cup \{i^k\},\ \mathbf{A}^k \leftarrow [\mathbf{A}^{k-1},\ \mathbf{A}_{i^k}]$
    $\mathbf{X}^k \leftarrow (\mathbf{A}^k)^\dagger \mathbf{Y},\ \hat{\mathbf{Y}}^k \leftarrow \mathbf{A}^k \mathbf{X}^k,\ \mathbf{R}^k \leftarrow \mathbf{Y} - \hat{\mathbf{Y}}^k$
end while
Output: $\mathbf{X}^* = \mathbf{X}^k$

The stopping criteria of the algorithm can be either a residual threshold or the sparsity level K. As mentioned previously, Algorithm 10 is mainly applicable to sources that emit a wide range of frequencies with the assumption that the corresponding signals occupy multiple frequencies at the frequency spectrum, such as harmonic signals. This assumption allows the algorithm to form groups of frequency bins and simultaneously attempt to identify the unknown DOAs. As empirically demonstrated in the following section, the grouping of frequencies proves to be quite beneficial, since the proposed method is able to suppress the undesirable effects of spatial aliasing.

In classical array processing, spatial aliasing refers to the ambiguities that arise when in the case of a ULA the sensor spacing is larger than half the wavelength of the incoming plane wave. From a compressed sensing point of view, this means that the DOA dictionary $\mathbf{A}$ will contain linear dependent columns and thus for certain support sets it will be impossible to distinguish between columns and identify the true DOAs. It is well known that MUSIC and other array processing methods suffer from the effects of spatial aliasing [13]. In classical beamforming, in order to avoid spatial aliasing, it is required that the sensor spacing is at most half the wavelength of the highest frequency. This can cause limitations in certain applications such as blind source separation (BSS), since for frequencies less than 16 kHz, the microphone spacing needs to be not larger than approximately 2 cm. For higher frequencies, sensor spacing requirements could be impossible to be satisfied.

Despite the fact that the proposed approach could suppress aliasing and overcome the above issues, the algorithm does not take into account possible sparsity or disjointness in the frequency domain and for this reason it is expected that its performance will degrade in these scenarios. To deal with these issues, further assumptions need to be made regarding the sources' energy distribution in the time-frequency domain.

5.6.3 Experimental results

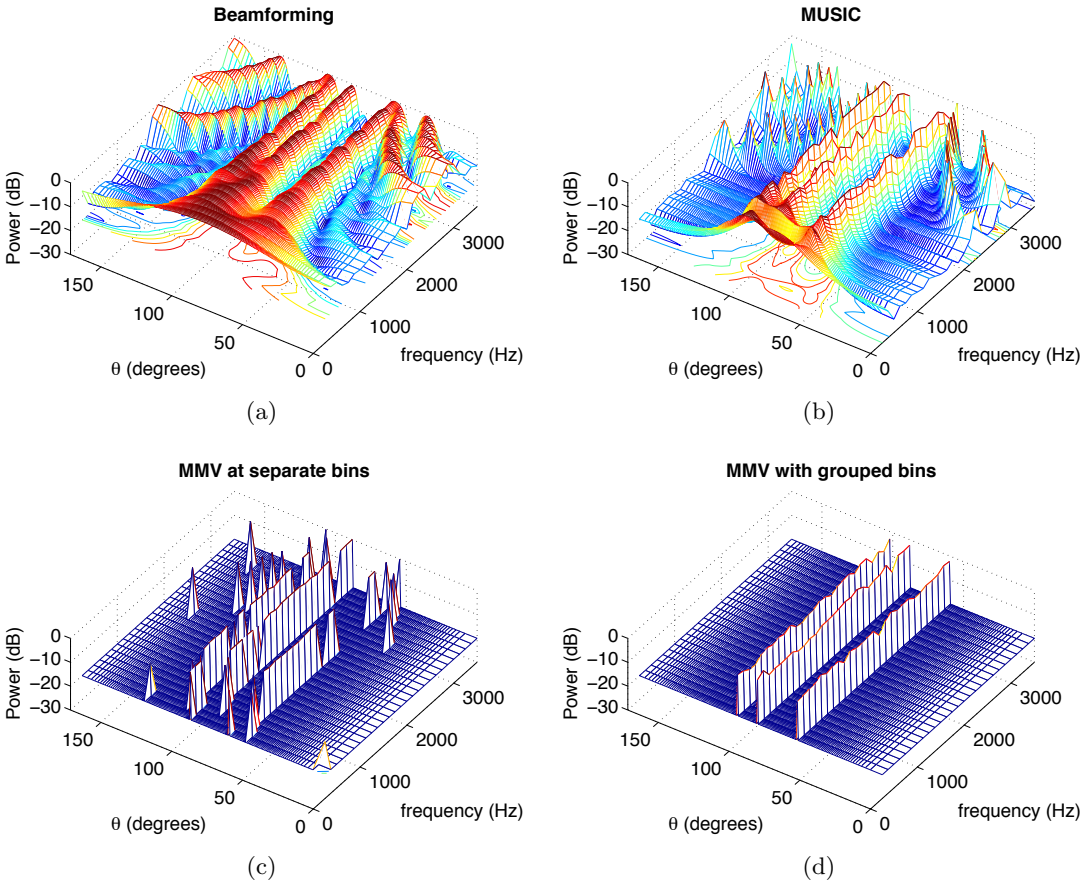


Figure 5.13: Spatial power spectrum estimate using (a) beamforming, (b) MUSIC, (c) MMV-PFP and (d) G-SOMP for $K = 3$ sources arriving on a ULA of $M = 8$ sensors with spacing set at half the wavelength corresponding to the middle frequency bin. The SNR is fixed at 10 dB.

This section presents some experimental results that highlight the advantages of the proposed approach over existing wideband array processing methods. Individual processing of each frequency bin under the sparse representations framework is also considered. In that case, the resulting MMV problem is solved at each frequency bin separately using the RA-PFP algorithm, which showed the best DOA estimation performance for narrow-

band ULAs. All of the following experiments consider a ULA of $M = 8$ sensors and the angular grid contains $N = 180$ potential DOAs.

For the first experiment, $K = 3$ wideband sources were considered, which emit waves with frequencies within the range of 250 Hz and 3 kHz and arrive on the array from different angles. The target sources were chosen to be located at 60°, 88° and 102° respectively. In that way, this experiment consists of two closely spaced sources and one that is well separated from the others. In order to partition the received signals into subbands, the data at each sensor were transformed using the STFT with non-overlapping frames. The sampling frequency was assumed to be 16 kHz and the duration of each time frame was chosen to be 8 ms. The additive white Gaussian noise (AWGN) at each sensor was set at SNR = 10 dB. Fig. 5.13 shows the resulting power spectrum estimates for all tested algorithms, namely classical beamforming, MUSIC, MMV-PFP and G-SOMP algorithms. The sensor spacing was set at half the wavelength of the middle frequency bin $f_c = (f_{max} - f_{min})/2 = 1375$ kHz. As can be seen, the methods of beamforming and MUSIC are both affected by spatial aliasing at the high frequencies above f_c. In addition, beamforming also shows its resoltuion limitations, as it fails to resolve the two closely spaced sources at all frequencies lower than f_c. Regarding the sparsity based methods, the proposed algorithm that considers grouped frequency bins displays the best performance showing three clear peaks at the true DOAs at all frequencies, outperforming all other methods.

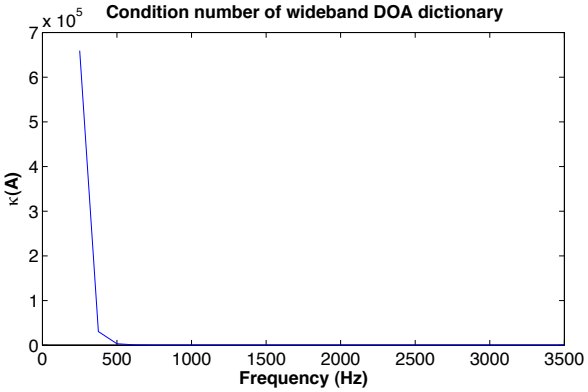


Figure 5.14: Condition number of wideband DOA dictionary as a function of the frequency.

In Fig. 5.13(c), it can be noted that the approach based on separate processing of each frequency bin introduces ambiguities; at high frequencies, which are possibly due to

spatial aliasing, and at low frequencies, which are due to algebraic aliasing, as discussed in [138]. This latter type of aliasing is concerned with ambiguities caused by using a narrowband ULA with geometry designed for a certain frequency at much lower frequencies. In that case, the resulting dictionary $\mathbf{A}_{f_{low}}$ becomes ill-conditioned and it turns out to be very hard to distinguish between subspaces, especially in the noisy case. Indeed, as can be seen in Fig. 5.14, where the condition number of the wideband DOA dictionary is plotted at all frequencies, the ratio of the largest to the minimum singular value of matrix $\mathbf{A}_f$ blows up at the very low frequencies. Both types of ambiguities are overcome by imposing the group sparsity constraint over all frequency bins.

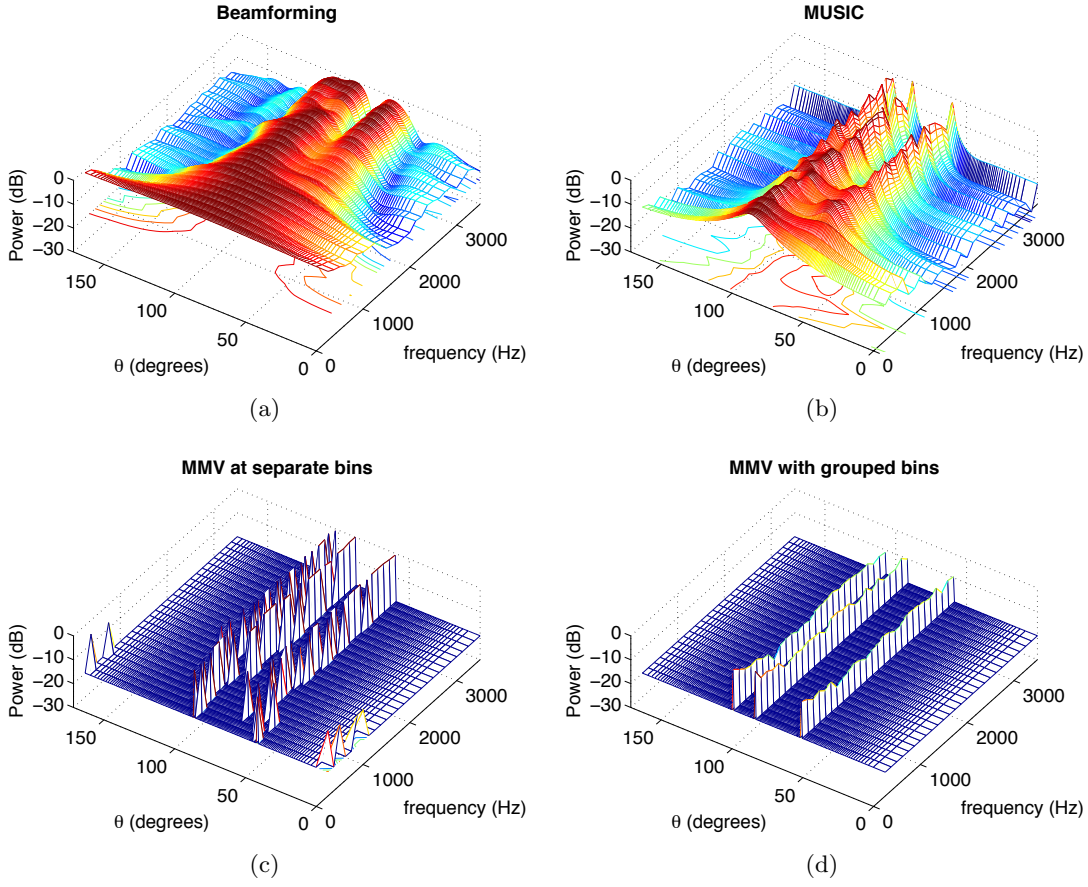


Figure 5.15: Spatial power spectrum estimate using (a) beamforming, (b) MUSIC, (c) MMV-PFP and (d) G-SOMP for $K = 3$ sources arriving on a ULA of $M = 8$ sensors with spacing set at half the wavelength corresponding to maximum frequency. The SNR is fixed at 10 dB.

Fig. 5.15 displays the spatial spectrum estimates for the same experimental setting, but using now a ULA with sensor spacing at half the wavelength of the highest frequency, which is the required spacing that eliminates the spatial aliasing for classic array

processing methods. Indeed, as can be seen, the effects of spatial aliasing at the high frequencies are not present anymore for both beamforming and MUSIC algorithms with the price of reduced resolution. As far as the sparsity based approaches are concerned, it can be noticed that although the resolution remains sharp, in fact both algorithms fail to identify the correct support set. This experiment demonstrates that the algebraic aliasing results in the deterioration of the performance of the sparsity based algorithms, and hence these methods are more sensitive to this type of aliasing rather than spatial aliasing, as shown in the previous experiment.

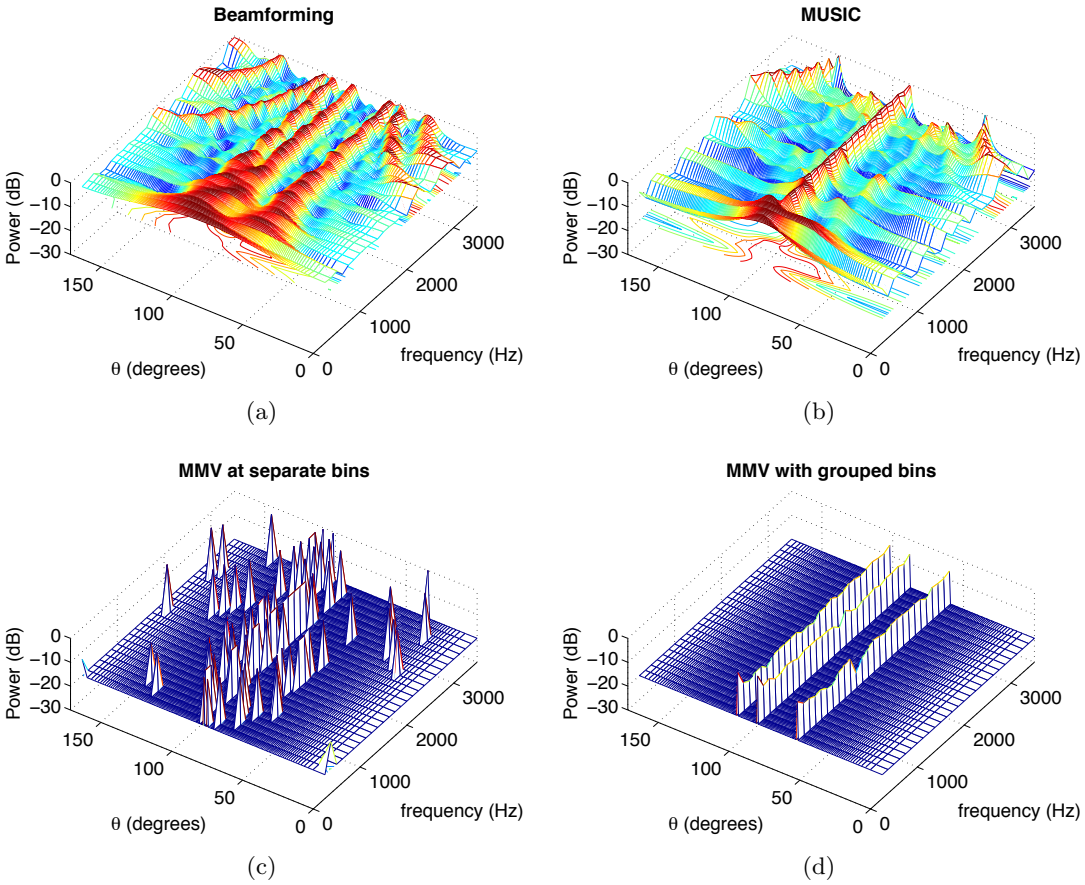


Figure 5.16: Spatial power spectrum estimate using (a) beamforming, (b) MUSIC, (c) MMV-PFP and (d) G-SOMP for $K = 3$ sources arriving on a ULA of $M = 8$ sensors with spacing set at half the wavelength corresponding to the middle frequency bin. Two of the incoming sources are highly correlated.The SNR is fixed at 0 dB.

Finally, simulation results are presented for the coherent source scenario. These results were accomplished by repeating the first experiment for a higher noise level such that SNR = 0 dB and letting two of the sources to be highly correlated to each other. Most notably, the MUSIC algorithm shows its limitations in that case, failing to resolve

all but one of the incoming sources (Fig. 5.16(b)). However, the G-SOMP algorithm succeeds in resolving all DOAs (Fig. 5.16(d)).

It has been shown how a structured sparse recovery approach can overcome limitations of conventional wideband DOA estimation methods such as spatial aliasing. Although these findings are based on empirical results and take advantage of the grouped frequency assumption, that might not always be met in practice, they show great potential for further research from both theoretical and practical perspective. The following section presents some further work that extends the specific wideband model to the near-field scenario, which is usually encountered in applications such as audio source localization in room environments.

5.6.4 Near-field wideband DOA estimation

In source separation, one is initially interested in localizing the sources before unmixing and eventually reconstructing them. When dealing with audio sources the far-field propagation model might not be very accurate considering that the sources lie near the microphone array. In that case, the wave equation yields a different solution and the plane wave approximation is no longer valid. For an omnidirectional point source in an unbounded space, the Green's function associated with the Helmholtz equation is given by a spherical wave emitted at $\mathbf{r}$ [139]:

$$\mathbf{g}(\mathbf{r}_m, \mathbf{r}, \omega) = \frac{\exp\{-j\frac{\omega}{c}\|\mathbf{r}_m - \mathbf{r}\|\}}{4\pi\|\mathbf{r}_m - \mathbf{r}\|} \tag{5.26}$$

where the Euclidean norm $\|\mathbf{r}_m - \mathbf{r}\|$ denotes the distance from the source point to the sensor placed at the point with coordinates $\mathbf{r}_m$.

Therefore, similarly to what presented earlier for the far-field case, one can construct a grid of potential locations $\mathbf{r}_i = [r_{x_i}, r_{y_i}, r_{z_i}]$ and use equation (5.26) to compute the impulse responses from the N grid points to each of the M sensors. It follows that in the case of wideband sources, several frequency bins need to be utilized, as described in Section 5.6.1.

To demonstrate the near-field wideband DOA estimation problem, the acoustic source localization problem was considered. A 1m $\times$ 1m $\times$ 3m rectangular room was assumed and $M = 4$ sensors were located at each corner. The resolution grid was chosen at 10 cm

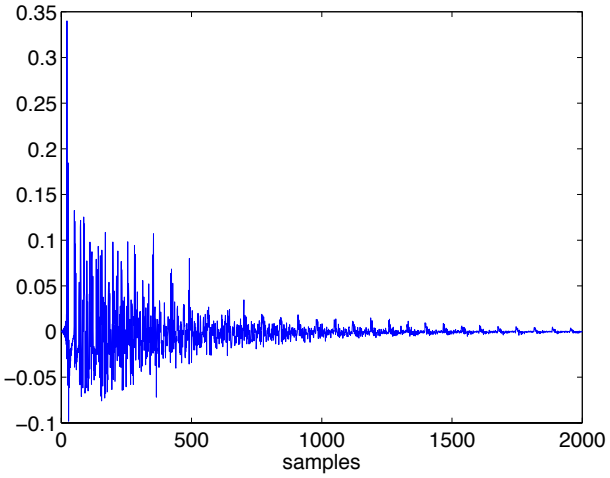
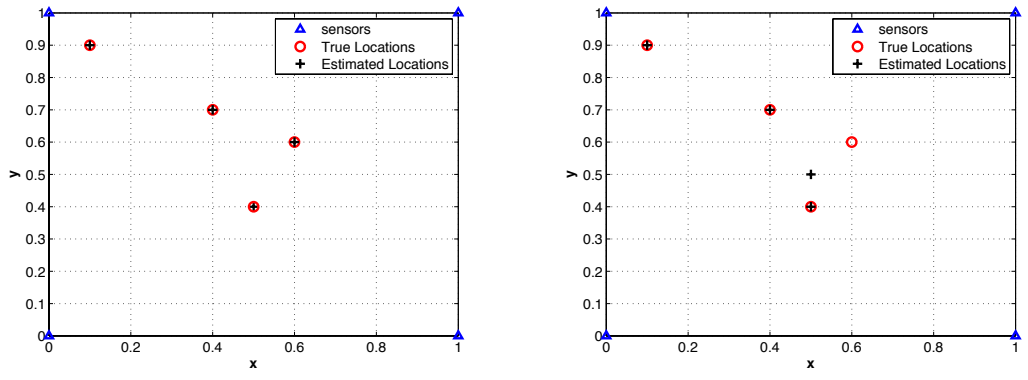


Figure 5.17: Impulse response for a simulated room environment with reverberation time $RT_{60} = 250$ msec.

resulting in a grid of $N = 81$. The number of sources was $K = 4$ and they were located at $[r_{x_1}, r_{y_1}, r_{z_1}] = [0.1, 0.9, 1.7]$, $[r_{x_2}, r_{y_2}, r_{z_2}] = [0.4, 0.7, 1.7]$, $[r_{x_3}, r_{y_3}, r_{z_3}] = [0.5, 0.4, 1.7]$ and $[r_{x_4}, r_{y_4}, r_{z_4}] = [0.6, 0.6, 1.7]$. For simplicity, it was assumed that the sources emit zero-mean white Gaussian signals, which are disjoint in the time domain and the sampling frequency was chosen to be $f_s = 8$ kHz. The impulse responses for each microphone were generated using the MATLAB package RIR generator developed by Emanuel A.P. Habets [140]. Fig. 5.17 displays an impulse response for the specific simulated room environment with reverberation time $RT_{60} = 250$ msec.



(a) (b)

Figure 5.18: Source localization of $K = 4$ sources in a simulated room with reverberation time of (a) $RT_{60} = 250$ msec and (b) $RT_{60} = 500$ msec.

The data at each sensor were obtained by superimposing the convolved signals with the corresponding filters. Next, the measured data were transformed into the time-

frequency domain using the STFT with non-overlapping segments of 16 msec. It was assumed that the sources utilize all considered frequency subbands. Subsequently, G-SOMP was used to localize the sources. The dictionary was generated by sampling the Green's function of equation (5.26) at the grid points for each frequency bin. Fig. 5.18 illustrates the localization results for two different reverberant conditions, 250 msec (2000 samples) and 500 msec (4000 samples). In the first case, the source localization is perfect, whereas in the second G-SOMP succeeds in localizing all but one source.

5.7 Discussion

This chapter has been concerned with the problem of DOA estimation and source localization in sensor arrays under the sparse representations framework. After presenting the ℓ_1-SVD method, which traces back to the innovative work of Malioutov et al. [3] and based on the recent theoretical results in the field of compressed sensing, it was initially shown that different sampling schemes of the impulse response of the array in consideration can lead to dictionaries with different properties. The presented analysis considered two different ways of discretizing the angular space in order to form a set of grid points. For each of those, there is a distinct column vector specified by the impulse response of the array. In more detail, from a compressed sensing point of view it was shown that discretizing the u-space, instead of the angular space, yields dictionaries with lower mutual coherence and better RIP properties.

Secondly, possible benefits in terms of the recovery performance of the jointly sparse inverse problem using NLA structures instead of typical ULAs were discussed. After applying the MMV-PFP algorithm, developed in Chapter 4, to the problem of DOA estimation, it was noticed that although the algorithm performs as good as the computationally expensive mixed norm minimization of the ℓ_1-SVD algorithm for well separated sources with respect to angular degrees, the algorithm's recoverability degrades in the scenario of closely spaced sources. To overcome this issue, the rank aware version of MMV-PFP algorithm was proposed. Experiments demonstrated that this approach can be generally beneficial, boosting the recovery performance when the dictionary exhibits high mutual coherence, while at the same time providing significant speed gains over ℓ_1-SVD due to its low complexity.

Next, the sparsity based DOA estimation problem was extended to the scenario of wideband sources. The observed data were transformed in the time-frequency domain and a "narrowband" approach was followed for each separate frequency bin. Moreover, it was shown that when the assumption that signals emit at the full frequency band holds, imposing the additional constraint of group sparsity over all or a subset of frequency bins can significantly improve the localization performance.

Finally, the problem of source localization of near-field wideband sources was considered and it was shown that by exploiting further assumptions (i.e. disjoint sources in the time-domain) localization of equal or more sources than the number of sensors is possible. However, the specific discussion was limited, considering only joint sources in the frequency domain. As stated in [141], the wideband near-field DOA estimation based on the sparse representations framework is a rather unexplored field and further research might lead to the development of more application oriented algorithms and methods.

Chapter 6

Off-grid DOA estimation with sparsity constraints

In Chapter 5, it was first shown how the DOA estimation problem can be formulated as a sparse recovery problem and then based on the CS framework, algorithms that can provide fast and efficient recovery under certain assumptions were developed. One of these assumptions requires that the sources arrive from directions that belong to a predefined discrete set of possible angles. However, in real applications this assumption might be very restrictive as in practice the unknown DOAs of the sources are continuous parameters i.e. $\theta \in [-\pi, \pi]$. Therefore, if the unknown DOAs do not belong in the angular grid the performance of the spatial sparsity methods will degrade due to errors caused by mismatches [142, 143].

A straightforward approach to minimize these model errors, caused by basis mismatches, is to increase the spatial resolution by making the grid finer and thus build up a dictionary with increased redundancy. However, the finer the angular grid the longer the recovery time needed by the CS algorithm. Besides, more sensors might be needed in that case, so that the model meets the CS requirement of incoherence for robust recovery, according to equations (5.10) and (5.18).

Assuming that the impinging sources are well separated (i.e. beyond the Rayleigh resolution limit) with respect to the given array aperture, it is expected that a finer grid will result in reducing the error related to algebraic ambiguities, which are due to

non exact matching of the true subspaces with the considered ones when the sources are off-grid. In that case, the trade-off between the spatial resolution and the convergence time poses the main limitation. In [3], the authors propose an adaptive grid refinement approach, in which the sparse recovery algorithm is applied iteratively starting from an initial coarse grid and successively making the grid finer around only the support obtained using the coarse grid and not universally. Although, this approach can significantly reduce the computational cost, it can still be quite expensive when high precision accuracy of DOA estimates is desired.

This chapter investigates the problem of off-grid DOA estimation, following a different route by treating the set of angular grid points as an additional unknown parameter. The proposed method is a fast iterative alternating descent algorithm, which improves the performance of the DOA estimation based on sparsity constraints, while keeping the computational complexity low.

6.1 Basis mismatch in DOA estimation

The problem of basis mismatch in CS and sparse representations has been addressed in several recent publications [142, 143, 144], as it happens to be a matter of concern in many applications. As discussed in [142], in compressive imaging the CS framework is often based on the assumption that the image of interest is sparse in the DFT basis. However, since this basis is a collection of atoms corresponding to a set of grid points of a parameter space, no matter how fine the specific grid is chosen to be, there is no guarantee that the signal of interest will be efficiently represented in the considered basis or dictionary. This implies that if the set or a subset of the frequencies that contain most or all of the energy in the signal do not exactly match with any of the grid points (e.g. a frequency falls in between two points of the grid set), then the CS reconstruction performance will degrade significantly. In the case of compressible signals, it might even be that the resulting representation is not sparse or compressible any more.

The DOA estimation problem is concerned with exact sparse representations, since it assumes that there are only few incoming sources on the array of sensors with regards to the size of the angular grid. If the choice of the angular grid is not appropriate, it might have disastrous effects in terms of localization of the sources. As can be seen in Fig. 6.1,

basis mismatches can result in complete disappearance of one of the impinging sources.

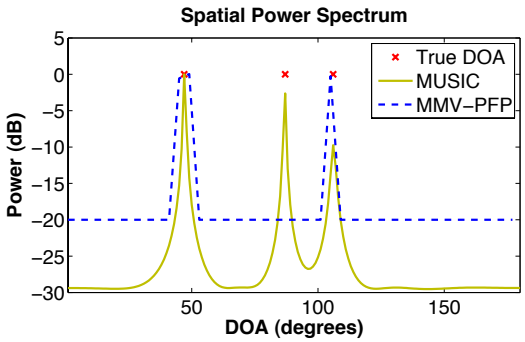


Figure 6.1: Spatial power spectrum estimate of MUSIC and MMV-PFP algorithms with basis mismatches due to off-grid DOAs.

Even though this might be the worst case scenario it is not guaranteed by any means that a finer grid can totally resolve such ambiguities and avoid this underestimation. However, a finer grid can still improve the DOA estimation performance in terms of RMSE with the price of an increased computational workload. As reported in [143], this is mainly due to the processing of a much larger dictionary.

6.2 Existing off-grid DOA estimation methods

Recently, Zhu et al. in [145, 146] addressed the problem of off-grid DOA estimation and developed a method to solve the total least squares problem with sparsity constraints. More specifically, this study considers the model:

$$\mathbf{y}(t) = [\mathbf{A} + \mathbf{E}]\mathbf{x}(t) + \mathbf{n}(t) \tag{6.1}$$

where $\mathbf{E}$ is a redundant matrix, which in the examined case represents the errors caused by potential mismatches. The authors propose the sparse regularized total least squares (SRTLS) algorithm in order to solve the problem:

$$\begin{aligned} &\min_{\mathbf{x}(t),\mathbf{n}(t),\mathbf{E}} \|[\mathbf{E}, \mathbf{n}(t)]\|_F^2 + \lambda\|\mathbf{x}(t)\|_1 \\ &\text{s.t.} \quad \mathbf{y}(t) = [\mathbf{A} + \mathbf{E}]\mathbf{x}(t) + \mathbf{n}(t). \end{aligned} \tag{6.2}$$

However, a thorough examination of the problem of equation (6.2) reveals that it requires the product of the unknown variables $\mathbf{E}$ and $\mathbf{x}$ to be optimized. In general, this is a non-convex optimization problem and for this reason a suboptimal method is proposed.

SRTLS in an iterative fashion alternates between estimates of $\mathbf{x}$ and $\mathbf{E}$. It first solves the Lasso problem for fixed $\mathbf{E}$ using the interior point solver SeDuMi and then for the estimated $\mathbf{x}$ it optimizes over $\mathbf{E}$:

$$\min_{\mathbf{E}} \|\mathbf{E}\|_F^2 + \|\mathbf{y}(t) - [\mathbf{A} + \mathbf{E}]\mathbf{x}(t)\|_2^2. \tag{6.3}$$

It can be shown that the above problem has a closed form solution [145]. Setting the derivative of (6.3) with respect to $\mathbf{E}$ to zero, it can be proven that the matrix $\mathbf{E}$ can be updated using the formula:

$$\mathbf{E} = [\mathbf{y}(t) - \mathbf{A}\mathbf{x}(t)]\mathbf{x}^H(t)[\mathbf{I} + \mathbf{x}(t)\mathbf{x}(t)]^{-1}. \tag{6.4}$$

The algorithm terminates when the difference between two consecutive estimates becomes smaller than a threshold and/or when the cost function falls under a chosen threshold. SRTLS consists a more general framework that can be used in any sparse reconstruction problem, for which the dictionary is approximately known and therefore it might introduce additional noise.

Another interesting approach to the problem of off-grid DOA estimation is the work in [147], which also addresses the problem of off-grid DOA estimation from a Bayesian perspective. The proposed sparse Bayesian inference (SBI) algorithm is an iterative algorithm, applicable in both cases of single and multiple snapshots. SBI enforces joint sparsity by imposing the same Laplacian prior over all measurements. The off-grid mismatches are modelled as a first order Taylor approximation problem. Other related work includes the Continuous Basis Pursuit algorithm [148], which makes use of a variety of approximation schemes, but it constrains the coefficient set to be nonnegative.

6.3 The proposed approach

In what follows, the SOMP-LS algorithm is proposed, which is an alternating descent algorithm in the vein of SRTLS. The algorithm uses Simultaneous Orthogonal Matching Pursuit at the first stage and then updates the dictionary with a least squares (LS) inversion.

The algorithm differs from SRTLS in how the mismatch errors are modelled. Fol-

lowing a similar approach with SBI, the problem of off-grid dictionary update is cast as an approximation problem. In that way, the developed method is a fast algorithm that exploits the parametric DOA dictionary to provide better estimation in the more general off-grid scenario.

6.3.1 Taylor approximation

Generally speaking, for any dictionary that is constructed by sampling a parametric function, as happens to be the case when it comes to DOA estimation, similar problems due to mismatches might arise.

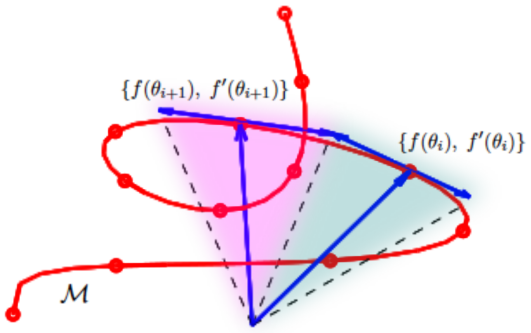


Figure 6.2: Illustration of approximation of the translational manifold $\mathcal{M}$ with a first order Taylor interpolation around the grid points (red dots).

Fig. 6.2 shows the nonlinear manifold $\mathcal{M}$ defined by the function $f(\theta)$ with $\theta \in [-\pi, \pi]$. The red dots correspond to the value of the function $f(\theta)$ at the selected grid points after the discretization of θ. Therefore, the set of vectors $f(\theta_i)$, $f(\theta_{i+1})$, etc. constitutes the overcomplete dictionary, which provide an approximation of this manifold [148].

In the off-grid case a single element of the manifold $\mathcal{M}$ will require the superimposition of several elements of the dictionary. One remedy to this problem is to augment the dictionary including interpolation functions that allow better approximation of the continuously shifted waveforms. It follows that if the function $f(\theta)$ and some of its derivatives are known at some point θ_i, then the function can be approximated at nearby

points using the Taylor series expansion:

$$f(\theta) \approx f(\theta_i) + (\theta - \theta_i) f'(\theta_i) + (\theta - \theta_i)^2 \frac{f''(\theta_i)}{2} + (\theta - \theta_i)^n \frac{f^{(n)}(\theta_i)}{n!} + \dots \tag{6.5}$$

The best linear Taylor approximation is given by the first order expansion, where the first two terms of the right hand side of the formula of equation (6.5) are only considered:

$$f(\theta) \approx f(\theta_i) + (\theta - \theta_i) f'(\theta_i) + \mathcal{O}((\theta - \theta_i)^2). \tag{6.6}$$

6.3.2 Problem formulation

Suppose now that the i-th plane wave impinges on the array from the angle $\tilde{\theta}_i$ that is not contained in the selected angular grid, namely $\tilde{\theta}_i \notin \{\theta_1, \dots, \theta_N\}$. In such a case, as described in [147], the corresponding vector $\mathbf{a}(\tilde{\theta}_i)$ for the off-grid DOA can be approximated by the first order Taylor expansion:

$$\mathbf{a}(\tilde{\theta}_i) \approx \mathbf{a}(\theta_i) + \mathbf{b}(\theta_i)(\tilde{\theta}_i - \theta_i) \tag{6.7}$$

where $\theta_i \in \{\theta_1, \dots, \theta_N\}$ is the nearest angle of the grid and $\mathbf{b}(\theta_i)$ is the first derivative of $\mathbf{a}(\theta_i)$ with respect to θ_i:

$$\mathbf{b}(\theta_i) = -j\pi \sin(\theta_i)\mathbf{p} \odot \mathbf{a}(\theta_i) \tag{6.8}$$

where $\mathbf{p} = \left[-\frac{M-1}{2},\ -\frac{M-3}{2},\ \dots,\ \frac{M-1}{2} \right]^T$ and $\circ$ denotes the element wise Hadamard product. It follows that one can define the redundant $M \times N$ matrix $\mathbf{B}$ with atoms $\mathbf{b}(\theta_i)$ for all N angles of the grid. The off-grid DOA model can then be formulated:

$$\mathbf{y}(t) = [\mathbf{A} + \mathbf{B}\Delta_\theta]\mathbf{x}(t) + \mathbf{n}(t) \tag{6.9}$$

with $\Delta_\theta = \text{diag}(\boldsymbol{\delta})$, $\boldsymbol{\delta} = [\delta_1, \dots, \delta_N]^T$ and $\delta_i = \tilde{\theta}_i - \theta_i$. In the above system of equations, both $\boldsymbol{\delta}$ and $\mathbf{x}(t)$ are unknowns and therefore after taking multiple time snapshots, the optimization problem can be formulated as:

$$\min_{\mathbf{X}, \boldsymbol{\delta}} \|\boldsymbol{\delta}\|_2^2 + \|\mathbf{Y} - [\mathbf{A} + \mathbf{B}\Delta_\theta]\mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_{1,2}. \tag{6.10}$$

However, the problem of equation (6.10), similar to (6.2) is non-convex and therefore it cannot be tackled using convex optimization. Thus, in the following, an alternating descent algorithm is proposed that iteratively shifts between estimates of $\mathbf{X}$ and $\boldsymbol{\delta}$ until the update of the specific matrices is no longer significant (e.g. falls below a predefined threshold).

6.3.3 Proposed alternating descent algorithm

A suboptimal way to solve the non-convex optimization problem in equation (6.10) is to reduce the problem to convex by minimizing over one parameter at a time. To do this, one can first look for a solution to the requlalized least squares problem by keeping the unknown vector $\boldsymbol{\delta}$ fixed and solve for $\mathbf{X}$. Therefore, at the k-th iteration, the algorithm needs to solve the MMV sparse recovery problem:

$$\min_{\mathbf{X}^k} \|\mathbf{X}^k\|_{1,2} \quad \text{s.t.} \quad \|\mathbf{Y} - [\mathbf{A} + \mathbf{B}\Delta_\theta^{k-1}]\mathbf{X}^k\|_F \leq \epsilon. \tag{6.11}$$

The sparse MMV problem (6.11) can be solved by mixed $\ell_{2,1}$ minimization or alternatively by greedy approaches such as SOMP or MMV-PFP, enforcing joint sparsity over the multiple vectors.

Once $\mathbf{X}$ has been updated, equation (6.10) is minimized over $\boldsymbol{\delta}$ keeping the current estimate of $\mathbf{X}$ fixed. In this case, the problem of equation (6.10) reduces to:

$$\min_{\boldsymbol{\delta}^k} \|\boldsymbol{\delta}^k\|_2^2 + \|\mathbf{Y} - [\mathbf{A} + \mathbf{B}\Delta_\theta^k]\mathbf{X}^k\|_F^2. \tag{6.12}$$

The problem (6.12) can be proven to have a closed form solution [145]. However, instead of taking the derivative and solving for $\boldsymbol{\delta}^k$, it is noted that for a single snapshot and as long as the cardinality of the support of $\mathbf{x}(t)$ is not larger than the number of sensors ($K < M$), the problem of (6.12) is equivalent to the least squares problem:

$$\mathbf{y}(t) - \mathbf{A}\mathbf{x}^k(t) = \mathbf{B}\Delta_{\mathbf{x}(t)}^k \boldsymbol{\delta}^k \tag{6.13}$$

where $\Delta_{\mathbf{x}(t)}^k = \text{diag}\big\{\big[x_1^k(t),\ \ldots,\ x_N^k(t)\big]\big\}$.

Considering now T time snapshots, it is straightforward to vectorize the resulting T

least squares problems:

$$\begin{bmatrix} \mathbf{y}(1) - \mathbf{A}\mathbf{x}^k(1) \\ \vdots \\ \mathbf{y}(T) - \mathbf{A}\mathbf{x}^k(T) \end{bmatrix} = \begin{bmatrix} \mathbf{B}\Delta^k_{\mathbf{x}(1)} \\ \vdots \\ \mathbf{B}\Delta^k_{\mathbf{x}(T)} \end{bmatrix} \boldsymbol{\delta}^k. \tag{6.14}$$

Therefore, at the k-th iteration the update to $\boldsymbol{\delta}$ will be:

$$\boldsymbol{\delta}^k = \mathbf{B}^{\dagger}_{\Delta^k_{\mathbf{X}}} \mathbf{R}^k \tag{6.15}$$

where $\mathbf{B}_{\Delta^k_{\mathbf{X}}} = \left[\mathbf{B}\Delta^k_{\mathbf{x}(1)}, \ldots, \mathbf{B}\Delta^k_{\mathbf{x}(T)}\right]^T$ and $\mathbf{R}^k = \mathbf{Y} - \mathbf{A}\mathbf{X}^k$.

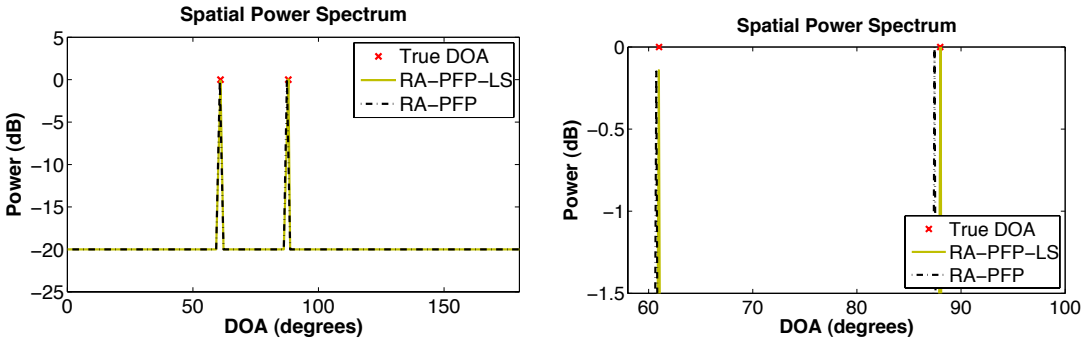
The proposed algorithm assumes that the sparsity level, namely the number K of the impinging on the ULA sources is known a priori. Therefore, after obtaining T time snapshots, the dimensionality of the MMV problem can be reduced by applying singular value decomposition (SVD) to the $M \times T$ measurement matrix $\mathbf{Y}$, as discussed in Chapter 5.

Algorithm 11 RA-PFP-LS Alternating Descent algorithm

1: Input: $\mathbf{A}$, $\mathbf{B}$, $\mathbf{Y}$, K, $l_{\max}$
2: Initialize: $k \leftarrow 0$, $\boldsymbol{\delta}^0 \leftarrow \mathbf{0}$,
3: **while** $k \leq l_{\max}$ **do**
4: $\quad k \leftarrow k + 1$
5: $\quad \mathbf{X}^k \leftarrow$ RA-PFP$([\mathbf{A} + \mathbf{B}\Delta^k_\theta], \mathbf{Y}, K)$
6: $\quad \mathbf{R}^k \leftarrow \mathbf{Y} - \mathbf{A}\mathbf{X}^k$, $\mathbf{B}_{\Delta^k_{\mathbf{X}}} = \left[\mathbf{B}\Delta^k_{\mathbf{x}(1)}, \ldots, \mathbf{B}\Delta^k_{\mathbf{x}(T)}\right]^T$
7: $\quad \boldsymbol{\delta}^k \leftarrow \mathbf{B}^{\dagger}_{\Delta^k_{\mathbf{X}}} \mathbf{R}^k$
8: $\quad$ **if** $\|\boldsymbol{\delta}^k - \boldsymbol{\delta}^{k-1}\| \leq \epsilon$ **then** exit; **end if**
9: **end while**
10: Output: $\mathbf{X}, \boldsymbol{\delta}$

The proposed alternating descent algorithm (Algorithm 11) is initialized with $\boldsymbol{\delta}^0 = \mathbf{0}$ and the K-term approximation to the problem (6.11) is obtained by running the RA-PFP algorithm for K iterations. Next, $\boldsymbol{\delta}$ is updated through equation (6.15). The algorithm iterates between these two steps and terminates when the difference between two consecutive updates of $\boldsymbol{\delta}$ falls below some chosen threshold. The final values of $\boldsymbol{\delta}$ provide an approximate estimate of the difference between the nearest $\theta_i \in \{\theta_1, \ldots, \theta_N\}$ and the true DOAs $\tilde{\theta}_i \notin \{\theta_1, \ldots, \theta_N\}$ for $i = 1, \ldots, K$. Fig. 6.3 illustrates an example of off-grid DOA estimation of $K = 2$ sources impinging on the array from the angles 61° and 88°, using an overcomplete dictionary of size 8×91. Although, the RA-PFP algorithm

yields the solution with minimum error, since the algorithm selects the dictionary atoms that correspond to the nearest angles, RA-PFP-LS further reduces this error by exploiting the Taylor approximation, providing better DOA estimates.



(a) (b)

Figure 6.3: Spatial spectrum estimate of $K = 2$ off-grid sources impinging on a ULA of $M = 8$ sensors using RA-PFP and RA-PFP-LS.

When compared to SRTLS, the proposed method replaces the Lasso solver at the regularization step with the greedy algorithm that allows for faster convergence especially in the case that multiple snapshots are considered. Therefore, the K-term approximation of RA-PFP provides faster convergence due to its algorithmic simplicity. It is evident that any greedy algorithm with joint sparsity structure could be used instead and the choice of RA-PFP is mainly based on the overall better performance that the algorithm displayed in the problem of DOA estimation, as seen in the previous chapter.

At the second step of dictionary update, the algorithm exploits the interpolation dictionary $\mathbf{B}$ and estimates the vector $\boldsymbol{\delta}$ of size $N \times 1$ instead of the $M \times N$ matrix $\mathbf{E}$. As shown in the following section, experiments favour the updating rule of equation (6.15) instead of the SRTLS update of (6.4).

6.4 Experimental Results

This section presents experimental results for the evaluation of the proposed off-grid DOA alternating descent algorithm. The algorithm is compared against the SBI algorithm and the ℓ_1-SVD algorithm, which assumes that K sources arrive from angles that exactly match K DOAs from the selected angular grid. For a fair comparison with the SRTLS approach, the RA-PFP-TLS algorithm is also derived, which under the same update rule

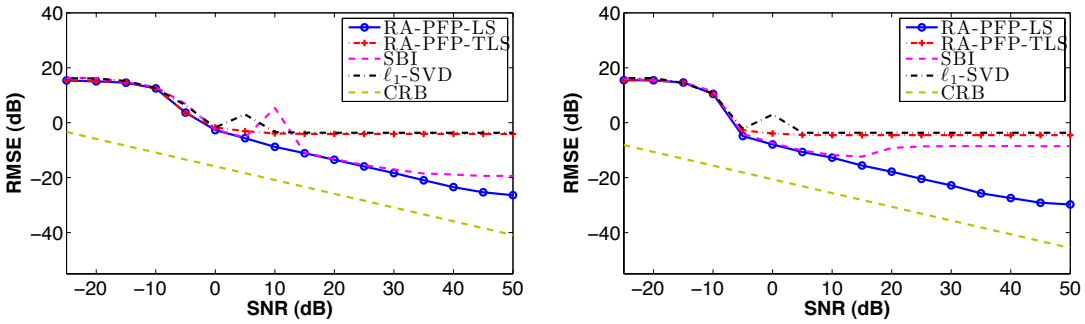
with the proposed method for $\mathbf{X}$ replaces equation (6.15) with:

$$\boldsymbol{\delta}^k = \text{diag}\Big\{\mathbf{B}^\dagger\big[\mathbf{Y} - \mathbf{A}\mathbf{X}^k\big]\big(\mathbf{X}^k\big)^H\big[\mathbf{I} + \mathbf{X}^k\mathbf{X}^k\big]^{-1}\Big\}. \tag{6.16}$$

In the following experiments, ULA structures of $M = 8$ and $M = 16$ sensors were considered and the u-space was uniformly discretized, resulting in a grid of $N = 91$ potential angles of arrival.

6.4.1 Sources with off-grid DOAs

For the first experiment, two zero mean narrowband far-field sources with equal power levels arriving on the ULA from directions 60.3° and 88.3° were considered and therefore the sparsity level was set at $K = 2$. The closest angles included in the grid were at 60.73° and 88.72°. The number of time snapshots was fixed at $T = 50$. For all tested algorithms, it was assumed that the sparsity level K is known a priori and the dimensionality of the measurements was reduced using the SVD method and thresholding the largest K singular values corresponding to the signal subspace. The additive noise at the sensors was white Gaussian and the noise level varied from -25 dB to 50 dB with a step size of 5 dB.



(a) (b)

Figure 6.4: Average RMSE of DOA estimation of $K = 2$ sources with off-grid directions (dB) vs SNR (dB) for (a) $M = 8$ and (b) $M = 16$ given $T = 50$ time snapshots.

Fig. 6.4 illustrates the average RMSE of the DOA estimation problem for all tested algorithms against the noise level for the two considered cases of ULA with $M = 8$ and $M = 16$ sensors. The results have been averaged over 100 trials. As expected, ℓ_1-SVD displays the worst performance with the largest error in most of the cases as the directions of the sources do not fall into the predefined angular grid. Among the off-grid

Table 6.1: Elapsed times (sec) of tested algorithms.

	$M = 8$		$M = 16$	
Alg/SNR	10dB	-10dB	10dB	-10dB
RA-PFP-LS	0.0321	0.0534	0.0216	0.0359
RA-PFP-TLS	0.0265	0.0316	0.0270	0.0447
SBI	0.5025	1.5335	0.2866	0.9246
ℓ_1-SVD	0.7501	0.6586	0.7849	0.7139

DOA estimation algorithms, the proposed method achieves the best overall performance. On the other hand, it can be seen that RA-PFP-TLS shows the poorest performance, not managing to achieve any significant performance gain over ℓ_1-SVD. This implies that the update rule of equation (6.15) results in better approximations than the one in (6.16). RA-PFP-LS also performs slightly better than the SBI algorithm in the case of $M = 8$ sensors especially for SNR values above 30dB (Fig. 6.4(a)). This performance gain increases for $M = 16$ sensors (Fig. 6.4(b)), as for some reason the RMSE of SBI algorithm suddenly degrades for SNR values higher than 15dB and eventually stalls at a much lower level when SNR becomes larger than 20dB. Although, one can argue that the first order Taylor interpolation might not provide the best approximation especially as the dimensionality of each column increases (i.e. by considering a larger number of sensors), it does not seem to be the case for RA-PFP-LS algorithm, since its performance remains intact. The dotted (mustard) line corresponds to the Cramér-Rao lower bound (CRB) for the specific DOA estimation problem, which is associated with the Fisher information and expresses a lower bound on the variance of any estimator. It is expected that the gap between the RA-PFP-LS algorithm curve and CRB line can be reduced if a higher order Taylor approximation is utilized. However, this will result in the inversion of a nonlinear least squares problem, which is computationally more demanding.

The average convergence time of each algorithm for two noise levels (10 dB and -10 dB) is shown in Table 6.1. In both cases examined ($M = 8$ and $M = 16$), the proposed approach along with RA-PFP-TLS was the fastest and when compared to SBI, RA-PFP–LS was at least 10 times faster.

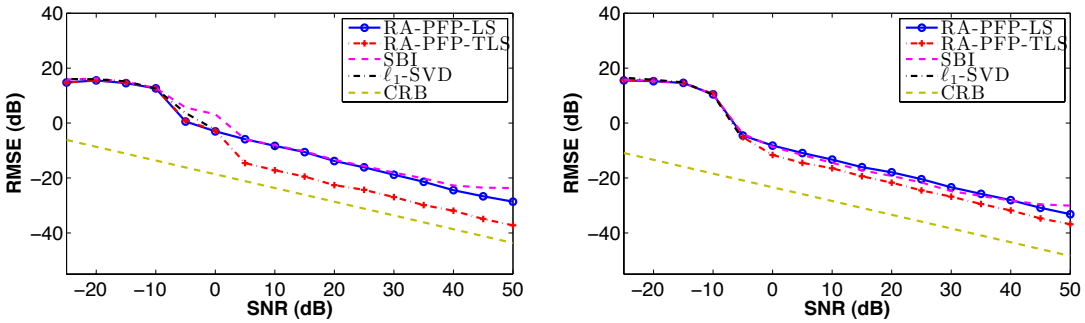
6.4.2 Sources with on-grid DOAs

The second experiment kept the same settings, but this time it was assumed that the $K = 2$ sources arrive on the ULA from the directions 60.73° and 88.72°, which are

Table 6.2: Elapsed times (sec) of tested algorithms.

	$M = 8$		$M = 16$	
Alg/SNR	10dB	-10dB	10dB	-10dB
RA-PFP-LS	0.0176	0.0560	0.0196	0.0499
RA-PFP-TLS	0.0133	0.0320	0.0198	0.0620
SBI	0.3797	1.5009	0.2647	1.4525
ℓ_1-SVD	0.7093	0.6613	0.8832	0.7872

included in the angular grid. Subsequently, the specific experiment attempted to examine the error introduced by the off-grid DOA estimation, when the set of DOAs of the sources is a subset of the discrete angular grid.



(a) (b)

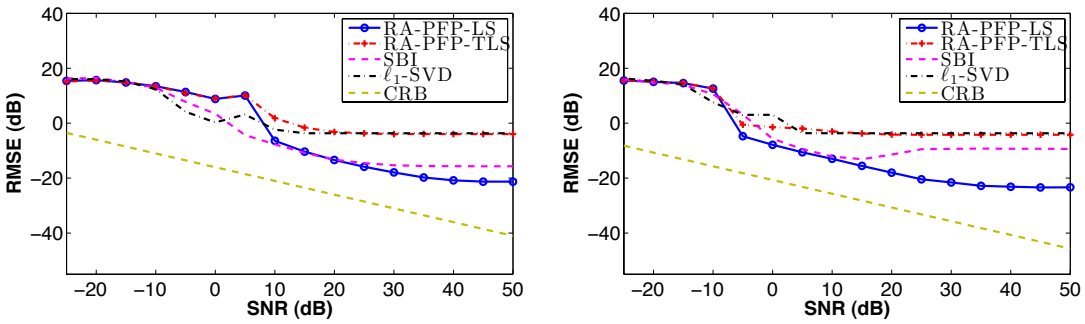
Figure 6.5: Average RMSE of DOA estimation of $K = 2$ sources with on-grid directions (dB) vs SNR (dB) for (a) $M = 8$ and (b) $M = 16$ given $T = 50$ time snapshots.

The simulation average results over 100 iterations are summarized in Fig. 6.5. As can be seen in Fig. 6.5(a), ℓ_1-SVD introduces error only for high noise levels (below 0 dB), but when the SNR is larger than 0 dB, it achieves perfect estimation of the unknown parameters. On the other hand, the off-grid DOA estimation methods introduce estimation error in all noise levels examined. Among them, RA-PFP-TLS performs better and closer to the CRB. This might be due to the underestimation, noticed in the off-grid case, which in the examined scenario proves to be beneficial. However, this performance gain reduces in the case when the number of sensors is $M = 16$ (Fig. 6.5(b)). Regarding the proposed RA-PFP-LS method and SBI algorithm, both perform closely in this experiment. It is worth noting that RA-PFP-LS performs similar to the case of the off-grid sources.

Table 6.2 summarizes the average convergence times for all algorithms. Once again RA-PFP-LS and RA-PFP-TLS provide the fastest convergence.

6.4.3 Off-grid DOA estimation with correlated sources

Next, experiments were carried out to evaluate the algorithms' performance in the scenario when the impinging sources are highly correlated. To do this, the first experiment was repeated, considering $K = 2$ highly correlated sources with directions 60.3° and 88.3°. The sources were generated as described in Section 5.5.4 of Chapter 5.



(a) (b)

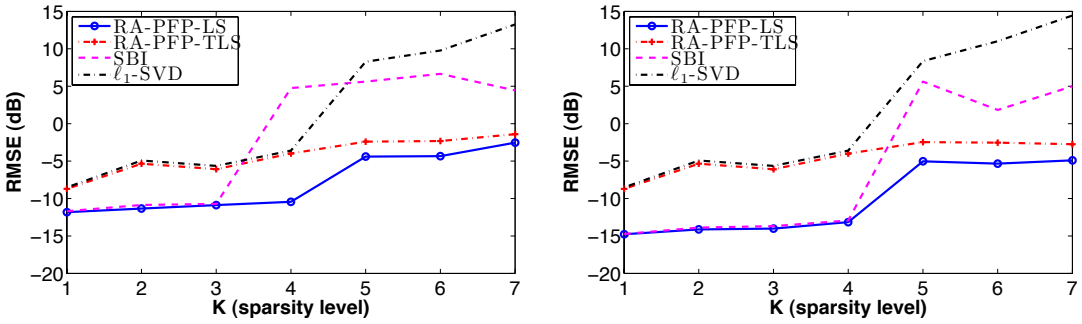
Figure 6.6: Average RMSE of DOA estimation of $K = 2$ correlated sources with off-grid directions (dB) vs SNR (dB) for (a) $M = 8$ and (b) $M = 16$ given $T = 50$ time snapshots.

While ℓ_1-SVD performance remains almost intact, the performance of the off-grid DOA estimation methods deteriorate, as shown in Fig. 6.6(a). RA-PFP-LS achieves the lowest RMSE values for low noise levels, but its performance degrades considerably when the noise is high (i.e. below 10dB). However, with the increase of the number of sensors, these ambiguities at the low SNR values seem to be resolved (Fig. 6.6(b)).

6.4.4 Variable number of off-grid sources

In the last experiment, an attempt was made to evaluate the off-grid DOA estimation as a function of the number of impinging sources, namely the sparsity level. For this reason, a ULA of $M = 8$ sensors was considered and the number of sources K varied from 1 to 7. In all cases, the sources arriving on the ULA were off-grid. The additive noise at the sensors was white Gaussian and the resulting SNR was fixed at 20 dB.

Fig. 6.7(a)-(b) show the RMSE as a function of the number of sources for $T = 50$ and $T = 200$ snapshots, respectively. The proposed method of RA-PFP-LS provides the best performance in both cases. More specifically, although its estimation performance is very close to the performance of the SBI algorithm for low sparsity levels, as the number of sources increases the RMSE also increases but not as much as for SBI.



(a) (b)

Figure 6.7: Average RMSE of DOA estimation (dB) vs the number of sources K with off-grid directions for (a) $T = 50$ and (b) $T = 200$ snapshots.

6.5 Discussion

This chapter considered the problem of off-grid DOA estimation under the sparse representation framework. After pointing out the undesirable effects that errors caused by basis mismatches can have in the problem of DOA estimation, a brief overview of the main existing approaches that tackle this problem was provided.

The proposed method is inspired by the SRTLS and SBI algorithms. Therefore, it attempts to instantaneously estimate the unknown support set, by identifying the appropriate subspace of the dictionary and update the dictionary by appropriately shifting its columns. However, this problem is non-convex and for this reason the proposed approach is based on a suboptimal alternating descent algorithm, which at the first stage attempts to identify the nearest directions of the sources to the ones included in the initial grid using the RA-PFP algorithm. It then updates the dictionary and the corresponding angular grid using a first order Taylor expansion. The process is repeated in an iterative fashion until the overall error has converged to some predefined threshold.

Experimental results have proven that RA-PFP-LS can overcome the resolution limitations of the standard sparsity model. When compared to other off-grid DOA methods such as the SBI algorithm, it also achieved slightly better performance, while in the coherent sources scenario the proposed algorithm outperformed all other tested methods. The simplicity of the greedy algorithm at the first stage combined with a single least squares inversion at the second stage of the algorithm resulted in the fastest convergence among the compared algorithms. More specifically, experiments showed that RA-PFP-LS converges at least 10 times faster than both SBI and ℓ_1-SVD, when two sources impinge on

an array of $M = 8$ sensors for all noise levels examined, regardless of whether the DOAs of the considered sources were off-grid or they were a subset of the initial angular grid.

Chapter 7

Conclusions & Future Work

This chapter concludes this thesis and summarizes its main contributions. Possible extensions for further work are also discussed.

7.1 Thesis summary

In this thesis, the general sparse recovery framework has been examined and its applicability to the problem of DOA estimation and source localization has been investigated.

For this purpose, after providing the background related to this thesis(Chapter 2), in Chapter 3, the problem of sparse representations and compressed sensing was studied from a practical point of view with the aim to develop fast recovery algorithms that exhibit good theoretical properties and are suitable for large scale problems. The Polytope Faces Pursuit algorithm was considered and its solution update strategy was modified appropriately, using the method of conjugate gradients. The resulting Conjugate Gradient Polytope Faces Pursuit algorithm proved to be efficient for large scale sparse recovery problems reducing the overall memory requirements of the original algorithm. Experimental and theoretical results demonstrated its robustness. The Stagewise Conjugate Gradient Polytope Faces Pursuit was also proposed, which further reduces the complexity, by allowing for several dictionary atoms to be added at each iteration.

In Chapter 4, the problem of structured sparsity was considered. Under the assumption that the dictionary coefficients form groups of atoms, the Group Polytope Faces Pursuit algorithm considering both $\ell_{2,1}$ and $\ell_{\infty,1}$ minimization problems was developed.

Although the theoretical results did not seem to favour a particular algorithm, empirical findings clearly showed the superiority of the GPFP algorithm based on the $\ell_{2,1}$ minimization criterion in the examined case, which only considered non-overlapping groups of equal size. Finally, the problem of joint sparsity that arises in the MMV sparse recovery context was investigated. Despite the fact that the worst case scenario theoretical results for the proposed MMV-PFP algorithm did not reveal any performance gain over the conventional sparsity algorithm, the experimental results indicated that the proposed algorithm provides improved recovery performance when the assumption of joint sparsity is met. In certain cases MMV-PFP also showed clear benefits over popular joint sparse recovery algorithms such as GOMP.

Chapter 5 thoroughly investigated the problem of DOA estimation under the spatial sparsity constraint, which makes the assumption that the number of sources to be localized is somewhat smaller than the number of grid points corresponding to the potential locations of the sources. The far-field and narrowband propagation model for uniform linear array structures was first considered. Adopting a compressed sensing approach, it was shown that the uniform discretization of the u-space can lead to dictionaries with better properties than the ones obtained by uniformly discretizing the angular space. Furthermore, the role of the geometry of the linear array to the sparse DOA recovery model was also examined. Although, the proposed MMV-PFP algorithm displayed its shortcomings in resolving closely spaced sources, the modified version of RA-PFP showed improved performance outperforming state-of-the-art sparsity based DOA estimation methods. Finally, the proposed framework was extended to the wideband and near-field scenarios, where it was shown how additional structure, if exploited, can result in the suppression of the undesirable effects of spatial aliasing.

In Chapter 6, the problem of off-grid DOA estimation was addressed. Considering that the assumption that the DOAs of the sources belong in a predefined set of grid points may not be realistic in practical situations, an alternating descent algorithm was developed, which at the first stage attempts to identify the nearest directions of the sources to the ones included in the initial grid using the RA-PFP algorithm and at the second stage updates the angular grid utilizing a first order Taylor expansion. Experimental results demonstrated the benefits of the proposed approach for the off-grid DOA

estimation problem.

7.2 Contributions

The principal contributions of this thesis are:

- the development of the stepwise and stagewise CG-PFP algorithms, as well as the analysis of the theoretical properties of the stepwise CG-PFP method (Chapter 3).
- the extension of the PFP algorithm to the group and joint sparsity scenarios (GPFP and MMV-PFP algorithms) and the investigation of their theoretical recovery guarantees (Chapter 4).
- the analysis of the DOA estimation problem based on the spatial sparsity assumption using tools from the field of compressed sensing (Chapter 5).
- the development of the modified version of MMV-PFP algorithm for the narrowband DOA estimation scenario (RA-PFP algorithm) and the proposal of a greedy method based on OMP in the wideband case (Chapter 5).
- the development of an alternating descent approach (RA-PFP-LS) to the problem of off-grid DOA estimation (Chapter 6).

7.3 Future work

Further work could follow on from the research of this thesis. This section suggests some possible new research avenues that have been identified throughout the work of this thesis.

Extensions of Polytope Faces Pursuit

Regarding the sparse recovery PFP algorithm several future directions could be possible. First of all, the investigation of the theoretical properties of the MMV-PFP algorithm, similar to the work presented in [119] for mixed $\ell_{2,1}$ norm minimization suggests one possibility for future research. A probabilistic average case analysis might explain the superior performance of MMV-PFP in the joint sparsity scenario, providing a formal description of the empirical findings presented in this thesis.

The work in [74] apart from the original PFP proposes an extension of the algorithm to the noisy scenario. Therefore, another potential route for further research would be the extension of the proposed PFP algorithms for conventional and structure sparsity to the noisy scenario. Investigation of the geometry and the theoretical properties of the noisy problem might provide better understanding of the problem and result in the development of efficient recovery methods.

Furthermore, one might consider the new alternative model of analysis sparsity [149], which has been proven to be of great potential especially in the field of compressed sensing [150]. Therefore, an attempt to explore the underlying geometry of the problem using the PFP method for the recovery of analysis sparsity inverse problems could bring new insight to this rather unexplored but promising area.

Extensions of the source localization sparse method

As discussed in Chapter 5, the problem of wideband source localization using arrays of sensors has not yet been fully explored. Although, in this thesis an effort was made to address several aspects of the problem, additional work could lead to further improvements. More specifically, investigating and deploying new more realistic assumptions regarding the structures in the frequency domain is one attractive direction. Furthermore, considering that this work focused only on point sources, the examination of the scenario of spatially spread signals, which happen to be the case in many physical situations, such as acoustic environments might serve as another promising research direction. This might pave the way for developing a spatial sparsity based scheme that tackles the problem of source localization in room environments.

Another consideration with potential for further improvements arises in the case of circular array geometries, which as discussed in [134] possess certain advantages over linear arrays. Since the focus of this thesis was on linear arrays, the extension of the results presented and the developed algorithms to the case of circular arrays is of great research interest for further work on this topic.

Extensions of the off-grid DOA estimation algorithm

The challenging problem of off-grid DOA estimation, which could also be viewed as a special case of the more general problem of basis mismatch in compressed sensing,

undoubtedly shows great potential for future research. With regard to the proposed sub-optimal alternating descent formulation of the original non-convex problem, a theoretical analysis has yet to be conducted and explained. Furthermore, the investigation of the proposed method's link with parametric dictionary learning (DL) approaches could be very beneficial, allowing further improvements in terms of accuracy in DOA estimation.

7.4 Closing remarks

The importance of sparse representations in signal processing is indisputable. Along with its relative field of compressed sensing, sparse representations have opened up new horizons for scientific research.

This thesis tackled the challenging problem of DOA estimation with arrays of sensors following a sparsity based approach, inspired by the recent advances in this field. Many issues were addressed resulting in the development of novel schemes for the problem of DOA estimation, that could also be of more general use in applications where spare inverse problems are encountered. Empirical findings indicate that the exploitation of the spatial sparsity assumption can overcome limitations of popular methods, such as the MUSIC algorithm in scenarios when the number of available samples is deficient, the sources are coherent or the noise level is too high. Furthermore, the algorithmic simplicity of greedy algorithms such as Polytope Faces Pursuit, especially when they incorporate additional structure, led to the development of low complexity DOA estimation algorithms that exhibit very fast convergence.

At the same time, this study gave rise to many other questions and highlighted critical points that could offer new possibilities for further research in the topic.

Bibliography

[1] M. D. Plumbley, "Recovery of sparse representations by polytope faces pursuit," in *Proceedings of the 6th International Conference on Independent Component Analysis and Blind Source Separation (ICA 2006)*, pp. 206–213, March 5–8 2006.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[3] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[4] A. Gretsistas, I. Damnjanovic, and M. D. Plumbley, "Gradient polytope faces pursuit for large scale sparse recovery problems," in *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2010)*, pp. 2030–2033, March 14–19 2010.

[5] A. Gretsistas and M. D. Plumbley, "Stagewise conjugate gradient polytope faces pursuit for large scale sparse recovery problems," in *Proceedings of the Conference on Information Representation and Estimation (INSPIRE 2010)*, (1 page), September 6–8 2010.

[6] A. Gretsistas and M. D. Plumbley, "Group polytope faces pursuit for recovery of block-sparse signals," in *Proceedings of the 10th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2012)*, pp. 255–262, March 12–15 2012.

[7] A. Gretsistas and M. D. Plumbley, "A multichannel spatial compressed sensing approach for direction-of-arrival estimation," in *Proceedings of the 9th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2010)*, pp. 458–465, September 27–30 2010.

[8] A. Gretsistas and M. D. Plumbley, "A source localization approach based on structured sparsity for broadband far-field sources," in *Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS11)*, p. 35, June 27–30 2011.

[9] A. Gretsistas and M. D. Plumbley, "An alternating descent algorithm for the off-grid DOA estimation problem with sparsity constraints," in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO 2012)*, pp. 874–878, August 27–31 2012.

[10] M. I. Skolnik, *Introduction to Radar Systems*. New York, NY, USA: McGraw-Hill, 1962.

[11] R. O. Nielsen, *Sonar Signal Processing*. Boston, MA, USA: Artech House, 1991.

[12] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*. Englewood Cliffs, NJ, USA: Prentice Hall, 1980.

[13] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[14] A. Hassanien, *Advanced Array Processing in the Presence of Complicated Spatio-Temporal Sources*. PhD thesis, McMaster University, 2005.

[15] I. Bilik, "Spatial compressive sensing for direction-of-arrival estimation of multiple sources using dynamic sensor arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 1754–1769, 2011.

[16] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624–638, 2006.

[17] S. F. Cotter, "Multiple snapshot matching pursuit for direction of arrival (DOA) estimation," in *Proceedings of the 15th European Signal Processing Conference (EUSIPCO 2007)*, pp. 247–251, September 3–7 2007.

[18] G. Tang and A. Nehorai, "Support recovery for source localization based on over-complete signal representation," in *Proceedings of the IEEE International Confer-*

ence on Acoustics Speech and Signal Processing (ICASSP 2010), pp. 2730–2733, March 14–19 2010.

[19] M. F. Duarte, "Localization and bearing estimation via structured sparsity models," in *Proceedings of the IEEE Statistical Signal Processing Workshop (SSP 2012)*, pp. 333–336, August 5–8 2012.

[20] C. Zheng, G. Li, H. Zhang, and X. Wang, "An approach of DOA estimation using noise subspace weighted ℓ_1 minimization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011)*, pp. 2856–2859, May 22–27 2011.

[21] H. L. Van Trees, *Optimum Array Processing.* New York, NY, USA: John Wiley & Sons, 2002.

[22] U. Madhow, *Fundamentals of Digital Communication.* Cambridge, UK: Cambridge University Press, 2008.

[23] M. S. Bartlett, "Smoothing periodograms from time series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, 1948.

[24] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[25] R. O. Schmidt, *A Signal Subspace Approach to Multiple Emitter Location Spectral Estimation.* PhD thesis, Stanford University, 1981.

[26] K. Lee and Y. Bresler, "Subspace-augmented MUSIC for joint sparse recovery with any rank," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2010)*, pp. 205–208, October 4–7 2010.

[27] I. Bilik, "Spatial compressive sensing approach for field directionality estimation," in *Proceedings of the IEEE Radar Conference*, pp. 1055–1059, May 4–8 2009.

[28] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way.* Burlington, MA, USA: Academic Press, 3rd ed., 2008.

[29] M. D. Plumbley, T. Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse representations in audio and music: From coding to source separation," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 995–1005, 2010.

[30] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[31] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 972–982, 2010.

[32] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 48–56, 2008.

[33] Y. Tsaig, *Sparse Solution of Underdetermined Linear Systems: Algorithms and Applications*. PhD thesis, Stanford University, 2007.

[34] B. L. Sturm, *Sparse Approximation and Atomic Decomposition: Considering Atom Interactions in Evaluating and Building Signal Representations*. PhD thesis, University of California, Santa Barbara, 2009.

[35] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[36] S. A. Abdallah and M. D. Plumbley, "Unsupervised analysis of polyphonic music by sparse coding," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 179–196, 2006.

[37] M. D. Plumbley, S. A. Abdallah, T. Blumensath, and M. E. Davies, "Sparse representations of polyphonic music," *Signal Processing*, vol. 86, no. 3, pp. 417–431, 2006.

[38] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, no. 11, pp. 2353–2362, 2001.

[39] R. Gribonval and S. Lesage, "A survey of sparse component analysis for blind source separation: Principles, perspectives, and new challenges," in *Proceedings*

of the 14th European Symposium on Artificial Neural Networks (ESANN 2006), pp. 323–330, April 26–28 2006.

[40] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[41] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.

[42] E. Ravelli, G. Richard, and L. Daudet, "Union of MDCT bases for audio coding," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1361–1372, 2008.

[43] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[44] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[45] M. H. Wright, "The interior-point revolution in optimization: History, recent developments, and lasting consequences," *Bulletin of the American Mathematical Society*, vol. 42, no. 1, pp. 39–56, 2004.

[46] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing.* New York, NY, USA: Springer, 2010.

[47] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[48] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 40–44, November 1–3 1993.

[49] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proceedings of the IEEE Computer Society Conference on Computer*

Vision and Pattern Recognition (CVPR 2006), vol. 1, pp. 895–900, June 17–22 2006.

[50] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006.

[51] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[52] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[53] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[54] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[55] M. Fornasier, *Theoretical Foundations and Numerical Methods for Sparse Recovery.* Berlin, Germany: Radon series on Computational and Applied Mathematics, De Gruyter, 2010.

[56] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2558–2567, 2002.

[57] A. Feuer and A. Nemirovski, "On sparse representation in pairs of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1579–1581, 2003.

[58] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.

[59] J.-J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1341–1344, 2004.

[60] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[61] B. Mulgrew, P. M. Grant, and J. S. Thompson, *Digital Signal Processing: Concepts and Applications*. Basingstoke, UK: Palgrave Macmillan, 1998.

[62] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[63] E. J. Candès, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 3, pp. 1433–1452, August 22–30 2006.

[64] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[65] E. J. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.

[66] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[67] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, UK: Cambridge University Press, 2012.

[68] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[69] M. F. Duarte, C. Hegde, V. Cevher, and R. G. Baraniuk, "Recovery of compressible signals in unions of subspaces," in *Proceedings of the 43rd Annual Conference on Information Sciences and Systems (CISS 2009)*, pp. 175–180, March 18–20 2009.

[70] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[71] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[72] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements," in *Proceedings of the 40th Annual Conference on Information Sciences and Systems (CISS 2006)*, pp. 207–212, March 22–24 2006.

[73] M. D. Plumbley, "On polar polytopes and the recovery of sparse representations," *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3188–3195, 2007.

[74] M. D. Plumbley, "Geometry and homotopy for ℓ_1 sparse representations," in *Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS'05)*, pp. 131–134, November 16–18 2005.

[75] A. Schrijver, *Theory of Linear and Integer Programming*. New York, NY, USA: John Wiley & Sons, 1998.

[76] M. D. Plumbley and M. Bevilacqua, "Sparse reconstruction for compressed sensing using stagewise polytope faces pursuit," in *Proceedings of the 16th International Conference on Digital Signal Processing (DSP 2009)*, pp. 14–21, July 5–7 2009.

[77] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins University Press, 1996.

[78] M. E. Davies and T. Blumensath, "Faster & greedier: Algorithms for sparse reconstruction of large datasets," in *Proceedings of the 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008)*, pp. 774–779, 2008.

[79] T. Blumensath and M. E. Davies, "Gradient pursuits," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2370–2382, 2008.

[80] T. Blumensath and M. E. Davies, "Stagewise weak gradient pursuits," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4333–4346, 2009.

[81] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2nd ed., 2003.

[82] J. R. Shewchuk, "*An Introduction to the Conjugate Gradient Method without the Agonizing Pain*." `http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf`, Carnegie Mellon University, 1994.

[83] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789–4812, 2008.

[84] J.-J. Fuchs, "Detection and estimation of superimposed signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1998)*, vol. 3, pp. 1649–1652, May 12–15 1998.

[85] D. L. Donoho and J. Tanner, "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1906, pp. 4273–4293, 2009.

[86] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.

[87] I. Damnjanovic, M. E. P. Davies, and M. D. Plumbley, "SMALLbox - An evaluation framework for sparse representations and dictionary learning algorithms," in *Proceedings of the 9th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2010)*, pp. 418–425, September 27–30 2010.

[88] I. Damnjanovic, M. E. P. Davies, and M. D. Plumbley, "*The SMALLbox, a Framework for Sparse Representation and Dictionary Learning*." `http://small-project.eu/software-data/smallbox`, (software).

[89] D. L. Donoho, V. Stodden, and Y. Tsaig, "*About SparseLab*," tech. rep., Stanford University, March 2007.

[90] V. Stodden, L. Carlin, D. L. Donoho, I. Drori, D. Dunson, M. Elad, S. Ji, J.-L. Starck, J. Tanner, V. Temlyakov, Y. Tsaig, and Y. Xue, "*SparseLab: Seeking Sparse*

Solutions to Linear Systems of Equations." `http://sparselab.stanford.edu/`, (software).

[91] R. D. N. M. Figueiredo and S. J. Wright, "*Gradient Projection for Sparse Reconstruction (GPSR).*" `http://www.lx.it.pt/~mtf/GPSR/`, (software).

[92] E. van den Berg, M. P. Friedlander, G. Hennenfent, F. Herrmann, R. Saab, and Ö. Yılmaz, "*SPARCO: A Testing Framework for Sparse Reconstruction*," tech. rep., University of British Columbia, October 2007.

[93] E. van den Berg, M. P. Friedlander, G. Hennenfent, F. Herrmann, R. Saab, and Ö. Yılmaz, "*SPARCO: A Toolbox for Testing Sparse Reconstruction Algorithms.*" `http://www.cs.ubc.ca/labs/scl/sparco/`, (software).

[94] T. Blumensath and M. E. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 4, pp. 1872–1882, 2009.

[95] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[96] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[97] L. He and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3488–3497, 2009.

[98] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[99] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 993–1009, 2009.

[100] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 101–111, 2003.

[101] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[102] F. Bach, "*Structured Sparsity-Inducing Norms through Submodular Functions.*" preprint, arXiv:1008.4220, `http://arxiv.org/pdf/1008.4220v3.pdf`, 2010.

[103] S. Sra, S. Nowozin, and S. J. Wright, *Optimization for Machine Learning.* Cambridge, MA, USA: MIT Press, 2011.

[104] Y. C. Eldar and M. Mishali, "Block sparsity and sampling over a union of subspaces," in *Proceedings of the 16th International Conference on Digital Signal Processing (DSP 2009)*, pp. 1–8, July 5–7 2009.

[105] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3075–3085, 2009.

[106] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[107] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3 - A Matlab software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 545–581, 1999.

[108] L. Meier, S. Van De Geer, and P. Bühlmann, "The group Lasso for logistic regression," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 70, no. 1, pp. 53–71, 2008.

[109] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2005.

[110] E. Berg and M. P. Friedlander, "*Joint-Sparse Recovery from Multiple Measurements*." preprint, arXiv:0904.2051, `http://arxiv.org/pdf/0904.2051v1.pdf`, 2009.

[111] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "A trust region algorithm for nonlinearly constrained optimization," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1152–1170, 1987.

[112] M. R. Celis, J. E. Dennis, and R. A. Tapia, "A trust region strategy for nonlinear equality constrained optimization," in *Proceedings of the SIAM Conference on Numerical Optimization*, pp. 71–82, June 12–14 1984.

[113] B. L. Sturm, M. G. Christensen, and R. Gribonval, "Cyclic pure greedy algorithms for recovering compressively sampled sparse signals," in *Proceedings of the 45th Asilomar Conference on Signals, Systems and Computers*, pp. 1143–1147, November 6–9 2011.

[114] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1996)*, vol. 3, pp. 1688–1691, May 7–10 1996.

[115] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: A recursive weighted minimum norm algorithm," *Electroencephalography and Clinical Neurophysiology*, vol. 95, no. 4, pp. 231–251, 1995.

[116] J. W. Phillips, R. M. Leahy, and J. C. Mosher, "MEG-based imaging of focal neuronal current sources," in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging*, vol. 3, pp. 1746–1750, October 21–28 1995.

[117] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proceedings of the 39th Asilomar Conference on Signals, Systems and Computers*, pp. 1537–1541, October 30–November 2 2005.

[118] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1135–1146, 2012.

[119] Y. C. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 505–519, 2010.

[120] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.

[121] R. Gribonval, H. Rauhut, K. Schnass, and P. Vandergheynst, "Atoms of all channels, unite! Average case analysis of multi-channel sparse recovery using greedy algorithms," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 655–687, 2008.

[122] V. Cevher, M. Duarte, and R. G. Baraniuk, "Distributed target localization via spatial sparsity," in *Proceedings of the 16th European Signal Processing Conference (EUSIPCO 2008)*, (5 pages), August 25–29 2008.

[123] A. C. Gurbuz, J. H. McClellan, and V. Cevher, "A compressive beamforming method," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, pp. 2617–2620, March 30–April 4 2008.

[124] V. Cevher, A. C. Gurbuz, J. H. McClellan, and R. Chellappa, "Compressive wireless arrays for bearing estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, pp. 2497–2500, March 30–April 4 2008.

[125] D. M. Malioutov, M. Cetin, J. W. Fisher III, and A. S. Willsky, "Superresolution source localization through data-adaptive regularization," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 194–198, August 4–6 2002.

[126] J. Zheng, M. Kaveh, and H. Tsuji, "Sparse spectral fitting for direction of arrival and power estimation," in *Proceedings of the IEEE/SP 15th Workshop on Statistical Signal Processing (SSP 2009)*, pp. 429–432, August 31–September 3 2009.

[127] G. Tzagkarakis, D. Milioris, and P. Tsakalides, "Multiple-measurement Bayesian compressed sensing using GSM priors for DOA estimation," in *Proceedings of*

the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2010), pp. 2610–2613, March 14–19 2010.

[128] A. C. Fannjiang, "*The MUSIC Algorithm for Sparse Objects: A Compressed Sensing Analysis.*" preprint, arXiv:1006.1678, `http://arxiv.org/pdf/1006.1678v4.pdf`, 2011.

[129] L. Carin, "On the relationship between compressive sensing and random sensor arrays," *IEEE Antennas and Propagation Magazine*, vol. 51, no. 5, pp. 72–81, 2009.

[130] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive MUSIC: Revisiting the link between compressive sensing and array signal processing," *IEEE Transactions on Information Theory*, vol. 58, no. 1, pp. 278–301, 2012.

[131] Z. Weng and X. Wang, "Support recovery in compressive sensing for estimation of direction-of-arrival," in *Proceedings of the 45th Asilomar Conference on Signals, Systems and Computers*, pp. 1491–1495, November 6–9 2011.

[132] K. Lee, Y. Bresler, and M. Junge, "Subspace methods for joint sparse recovery," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3613–3641, 2012.

[133] E. J. Candès and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.

[134] J.-J. Fuchs, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 702–709, 2001.

[135] K. Varma, *Time-Delay-Estimate Based Direction-of-Arrival Estimation for Speech in Reverberant Environments.* PhD thesis, Virginia Polytechnic Institute and State University, 2002.

[136] N. Mitianoudis and M. E. Davies, "Using beamforming in the audio source separation problem," in *Proceedings of the 7th International Symposium on Signal Processing and Its Applications (ISSPA 2003)*, vol. 2, pp. 89–92, July 1–4 2003.

[137] T. Yoshioka, T. Nakatani, M. Miyoshi, and H. G. Okuno, "Blind separation and dereverberation of speech mixtures by joint optimization," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 69–84, 2011.

[138] Z. Tang, G. Blacquiere, and G. Leus, "Aliasing-free wideband beamforming using sparse signal representation," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3464–3469, 2011.

[139] E. Habets, "*Room Impulse Response (RIR) Generator*," tech. rep., Univeristy of Erlangen-Nuremberg, September 2010.

[140] E. Habets, "*Room Impulse Response (RIR) Generator*." `http://home.tiscali.nl/ehabets/rir_generator.html`, (software).

[141] P. T. Boufounos, P. Smaragdis, and B. Raj, "Joint sparsity models for wideband array processing," in *Proceedings of SPIE: Wavelets and Sparsity XIV*, vol. 8138, 81380K, September 27 2011.

[142] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.

[143] L. L. Scharf, E. K. P. Chong, A. Pezeshki, and J. R. Luo, "Sensitivity considerations in compressed sensing," in *Proceedings of the 45th Asilomar Conference on Signals, Systems and Computers*, pp. 744–748, November 6–9 2011.

[144] D. H. Chae, P. Sadeghi, and R. A. Kennedy, "Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals," in *Proceedings of the 2nd International Conference on Future Computer and Communication (ICFCC 2010)*, vol. 2, pp. 739–743, May 21–24 2010.

[145] H. Zhu, G. Leus, and G. B. Giannakis, "Sparse regularized total least squares for sensing applications," in *Proceedings of the 11th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2010)*, pp. 1–5, June 20–23 2010.

[146] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2002–2016, 2011.

[147] Z. Yang, L. Xie, and C. Zhang, "*Off-Grid Direction of Arrival Estimation Using Sparse Bayesian Inference*." preprint, arXiv:1108.5838, `http://arxiv.org/pdf/1108.5838v4.pdf`, 2011.

[148] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4735–4744, 2011.

[149] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, 2007.

[150] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

Appendix A

Vector & matrix norms

A.1 Vector norms

Given a vector space V, then a norm, denoted by $\|\cdot\|$, is defined as a function $f\colon V \to \mathbb{R}$ which satisfies the following properties:

(i) $f(\mathbf{x}) \geq 0 \quad \mathbf{x} \in V$

(ii) $f(\mathbf{x}) = 0 \quad$ if and only if $\mathbf{x} = \mathbf{0}$

(iii) $f(a\mathbf{x}) = |a| f(\mathbf{x}) \quad a \in \mathbb{R}, \mathbf{x} \in V$

(iv) $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \quad \mathbf{x}, \mathbf{y} \in V.$

Therefore, a norm f is a mapping of the vector space onto $\mathbb{R}$ and designates a measure of the size of the vector $\mathbf{x}$. Properties (i) and (ii) require that the size of the norm is nonnegative, property (iii) known also as positive homogeneity, requires its size to be scaled as the vector is scaled and property (iv) is the triangle inequality.

Consider the N-dimensional space $V = \mathbb{R}^N$. An obvious class of vector norms for $\mathbb{R}^N$ are the ℓ_p-norms defined by:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p}, \quad p \geq 1 \tag{A.1}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$. Among the ℓ_p norms, ℓ_1, ℓ_2 and ℓ_∞ are the most important:

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_N| \tag{A.2a}$$

$$\|\mathbf{x}\|_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_N|^2)^{1/2} \tag{A.2b}$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq N} |x_i|. \tag{A.2c}$$

A.2 Properties of vector norms

An immediate consequence of the triangle inequality is that for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, it also holds:

$$\|\mathbf{x} - \mathbf{y}\| \geq \|x\| - \|\mathbf{y}\|. \tag{A.3}$$

Another interesting property concerning ℓ_p-norms is the Cauchy-Schwarz inequality:

$$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \tag{A.4}$$

which is a special case of the more general Holder inequality:

$$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{A.5}$$

All ℓ_p norms on $\mathbb{R}^N$ are equivalent. In particular, it holds that:

$$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{N}\|\mathbf{x}\|_2 \tag{A.6a}$$

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{N}\|\mathbf{x}\|_\infty \tag{A.6b}$$

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq N\|\mathbf{x}\|_\infty. \tag{A.6c}$$

A.3 Matrix norms

As with the vector norms, a matrix norm on the $\mathbb{R}^{M \times N}$ vector space is defined as a function f: $\mathbb{R}^{M \times N} \to \mathbb{R}$, for which the following properties hold:

(i) $f(\mathbf{A}) \geq 0 \quad \mathbf{A} \in \mathbb{R}^{M \times N}$

(ii) $f(\mathbf{A}) = 0 \quad$ if and only if $\mathbf{A} = \mathbf{0}_{M \times N}$

(iii) $f(a\mathbf{A}) = |a| f(\mathbf{A}) \quad a \in \mathbb{R},\ \mathbf{A} \in \mathbb{R}^{M\times N}$

(iv) $f(\mathbf{A}+\mathbf{B}) \leq f(\mathbf{A}) + f(\mathbf{B}) \quad \mathbf{A}, \mathbf{B} \in \mathbb{R}^{M\times N}.$

The most frequently used matrix norms in linear algebra are the Frobenius norm, defined by:

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} |a_{ij}|^2} = \sqrt{\mathrm{tr}(\mathbf{A}^T\mathbf{A})} = \sqrt{\sum_{i=1}^{\min\{M,N\}} \sigma_i^2} \tag{A.7}$$

where σ_i represents the i-th singular value of $\mathbf{A}$, and the induced ℓ_p norms defined by:

$$\|\mathbf{A}\|_p = \max_{\mathbf{x}\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}. \tag{A.8}$$

Therefore, for $p = 1$ it is:

$$\|\mathbf{A}\|_1 = \max_{0\leq j\leq N} \sum_{i=1}^{M} |a_{ij}| \tag{A.9}$$

while for $p = \infty$:

$$\|\mathbf{A}\|_\infty = \max_{0\leq i\leq M} \sum_{j=1}^{N} |a_{ij}|. \tag{A.10}$$

In the special case that the induced norm is the $p = 2$ norm, known also as the spectral norm, it holds:

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})} = \sigma_{\max}(\mathbf{A}) \tag{A.11}$$

where $\sigma_{\max}$ denotes the maximum singular value of $\mathbf{A}$ and $\lambda_{\max}$ the maximum eigenvalue of $\mathbf{A}^T\mathbf{A}$.

A.4 Properties of matrix norms

For the spectral and Frobenius norms it holds that:

$$\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \leq \sqrt{r}\|\mathbf{A}\|_2 \tag{A.12}$$

where $r = \mathrm{rank}(\mathbf{A})$. Other equivalence properties include:

$$\frac{1}{\sqrt{N}}\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{M}\|\mathbf{A}\|_\infty \tag{A.13a}$$

$$\frac{1}{\sqrt{M}}\|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{N}\|\mathbf{A}\|_1. \tag{A.13b}$$

Appendix B

The Cramér-Rao bound for DOA estimation

In estimation theory, the Cramér-Rao bound (CRB) provides the lowest bound on the covariance matrix of any unbiased estimator of the parameter $\boldsymbol{\theta}$. More precisely, CRB states that for the covariance matrix of an unbiased estimator $\hat{\boldsymbol{\theta}}$ it holds that:

$$\mathrm{C}(\hat{\boldsymbol{\theta}}) \triangleq \mathrm{E}\Big\{ \big[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\big] \big[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\big]^T \Big\} \geq \mathrm{C}_{CRB}(\hat{\boldsymbol{\theta}}) \triangleq \mathbf{J}^{-1} \tag{B.1}$$

where $\mathbf{J}$ represents the Fisher information matrix [21]. The Fisher information is a measure of the amount of information that an observed random variable $\mathbf{y}$ contains about the unknown parameter $\boldsymbol{\theta}$ upon which the probability of $\mathbf{y}$ depends. It is defined by:

$$\mathbf{J} = \mathrm{E}\left\{ \left[\frac{\partial \log p_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T \left[\frac{\partial \log p_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \right\} = -\mathrm{E}\left\{ \left[\frac{\partial^2 \log p_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}} \right] \right\} \tag{B.2}$$

where $p_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})$ is the conditional on $\boldsymbol{\theta}$ probability density function of the random variable $\mathbf{y}$ or equivalently the likelihood function for $\boldsymbol{\theta}$. Any unbiased estimator that achieves the CRB of inequality (B.1) and therefore yields the lowest possible RMSE, is a minimum variance unbiased estimator. Such an estimator is usually referred to as an efficient estimator.

For the DOA estimation problem using arrays of sensors, the following model is

considered:

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) + \mathbf{n}(t) \tag{B.3}$$

where $\mathbf{y}(t)$ is the M-dimensional observations vector, $\boldsymbol{\theta} \in \mathbb{R}^K$ is the vector containing the unknown parameters corresponding to the DOAs of the incoming K signals $\mathbf{x}(t)$ and $\mathbf{n}(t)$ is the additive at the sensors zero-mean Gaussian noise i.e. $\mathbf{n} \sim \mathcal{N}(0, \sigma^2)$. Therefore, the observations covariance matrix is given by:

$$\mathbf{R}_\mathbf{y} \triangleq \mathrm{E}\{\mathbf{y}(t)\mathbf{y}^H(t)\} = \mathbf{A}(\boldsymbol{\theta})\mathrm{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\}\mathbf{A}^H(\boldsymbol{\theta}) + \mathrm{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \mathbf{A}(\boldsymbol{\theta})\mathbf{R}_\mathbf{x}\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2\mathbf{I}. \tag{B.4}$$

According to [21], the elements of the Fisher information matrix for the specific model are given by:

$$\mathbf{J}_{ij} = \mathrm{tr}\left[\mathbf{R}_\mathbf{y}^{-1}\frac{\partial \mathbf{R}_\mathbf{y}}{\partial \theta_i}\mathbf{R}_\mathbf{y}^{-1}\frac{\partial \mathbf{R}_\mathbf{y}}{\partial \theta_j}\right]. \tag{B.5}$$

Combining equations (B.4) and (B.5) and after further manipulation, the CRB for the DOA estimation problem is calculated and it is given by:

$$\mathrm{C}_{CRB}(\hat{\boldsymbol{\theta}}) = \frac{\sigma^2}{2T_s}\left\{\mathrm{Re}\Big[\big[\mathbf{R}_\mathbf{x}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{R}_\mathbf{y}\mathbf{A}(\boldsymbol{\theta})\mathbf{R}_\mathbf{x}\big] \odot \big[\mathbf{B}^H(\boldsymbol{\theta})P^\perp_{\mathbf{A}(\boldsymbol{\theta})}\mathbf{B}(\boldsymbol{\theta})\big]\Big]\right\}. \tag{B.6}$$

where T_s is the number of snapshots taken, $P^\perp_{\mathbf{A}(\boldsymbol{\theta})}$ is the orthogonal projection onto the noise subspace and $\mathbf{B}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times K}$ is the matrix given by:

$$\mathbf{B}(\boldsymbol{\theta}) = \left[\frac{\partial \mathbf{a}(\theta_1)}{\partial(\theta_1)}, \frac{\partial \mathbf{a}(\theta_2)}{\partial(\theta_2)}, \ldots, \frac{\partial \mathbf{a}(\theta_K)}{\partial(\theta_K)}\right]. \tag{B.7}$$

When the incoming signals are uncorrelated, the sources are not too closely spaced and the covariance matrix $\mathbf{R}_\mathbf{x}$ is not close to singular, then the CRB for the DOA estimation problem of equation (B.6) can be well approximated by:

$$\mathrm{C}_{CRB}(\hat{\boldsymbol{\theta}}) \cong \frac{\sigma^2}{2T_s}\left\{\mathrm{Re}\Big[\mathbf{R}_\mathbf{x} \odot \big[\mathbf{B}^H(\boldsymbol{\theta})P^\perp_{\mathbf{A}(\boldsymbol{\theta})}\mathbf{B}(\boldsymbol{\theta})\big]\Big]\right\}. \tag{B.8}$$